Unified Cybersecurity Platform Proposal

Synthesis of 21 AI Security Projects from girdav01

Document Version: 1.0
Date: February 12, 2026
Author: DeepAgent Analysis

Table of Contents

---

1. Executive Summary

Overview

This document presents a comprehensive analysis of 21 cybersecurity projects developed by girdav01, all focused on AI/ML security and modern cyber defense. These projects collectively represent a pioneering effort to address the emerging security challenges posed by artificial intelligence systems—from protecting AI infrastructure to detecting AI-enabled threats.

Project Portfolio Summary

Metric	Value
Total Projects	21
Primary Language	Python (16 projects)
Secondary Languages	TypeScript, JavaScript, Go, Shell
Time Period	November 2025 - February 2026
Core Focus	AI/ML Security, Shadow AI Detection, Agentic AI Security

Key Themes Identified

1. **Shadow AI Detection & Control** - Multiple tools for discovering unauthorized AI deployments
2. **Agentic AI Security** - Detection and response for autonomous AI systems
3. **AI Supply Chain Security** - Protecting ML pipelines and model integrity

4. **GenAI Attack Prevention** - Defending against AI-powered threats (phishing, deepfakes)
5. **Security Training & Education** - Hands-on OWASP LLM Top 10 training
6. **SOC Automation** - AI-powered security operations
7. **Threat Intelligence** - AI-focused TI feeds and telemetry standards

Strategic Value

These 21 projects, when unified, form the foundation for a **comprehensive AI-Native Cybersecurity Platform** that addresses the full spectrum of AI security challenges—from protecting AI systems to leveraging AI for defense, and from compliance to active threat hunting.

---

2. Project Analysis

2.1 Categorization by Functional Domain

🔍 DETECTION DOMAIN (7 Projects)

Project	Purpose	Key Capabilities
AIDisco	Shadow AI Scanner	Cross-platform LLM detection, Docker/WSL2 scanning, SIGMA rules
V1ShadowAI	Vision One Shadow AI	Native V1 integration for unauthorized AI detection
AICrawler	AI Service Discovery	Multi-agent crawler, risk scoring, detection rule generation
LieDetector	Social Engineering Detection	Deepfake/fraud detection, trust scoring, behavioral analysis
AntiphishingGenAI	GenAI Phishing Detection	Linguistic analysis, AI content detection, attachment scanning
AIDataGuard	Data Security Monitoring	File integrity, log inspection, endpoint management
AITelemetry	Security Telemetry	OCSF schema, RFC standard, cross-platform collectors

🛡️ PREVENTION DOMAIN (5 Projects)

Project	Purpose	Key Capabilities
AISupplyChain	Supply Chain Security	AI-BOM, artifact signing, lineage tracking, CoSAI controls
AgenticAIDR	Agentic AI D&R	Real-time action tracking, policy enforcement, kill-switch
AISEC	Security Framework	Comprehensive AI security framework
UniversalGuardrail	Guardrail Standard	Universal API for AI guardrails
AIGuardAPIDemo	Guard API Demo	Demonstration of AI Guard capabilities

🎯 SECURITY TESTING DOMAIN (3 Projects)

Project	Purpose	Key Capabilities
IndirectPromptTester	Prompt Injection Testing	118 attack vectors, file generation, difficulty grading
AIHoneypot	Vulnerable AI Demos	OWASP LLM Top 10 demos, honeypot deployment
AISECTraining	Security Training	Hands-on workshops, dual security approaches

🤖 SOC AUTOMATION DOMAIN (3 Projects)

Project	Purpose	Key Capabilities
AutomatedSOC	Agentic SOC Triage	LLM-powered triage, risk prioritization, response actions
cyberAgents	Multi-Agent Analysis	9 specialist agents, threat intelligence integration
VisionOneSkills	Endpoint Skills	7 production skills, MITRE mapping, cross-platform

🔗 INTEGRATION DOMAIN (2 Projects)

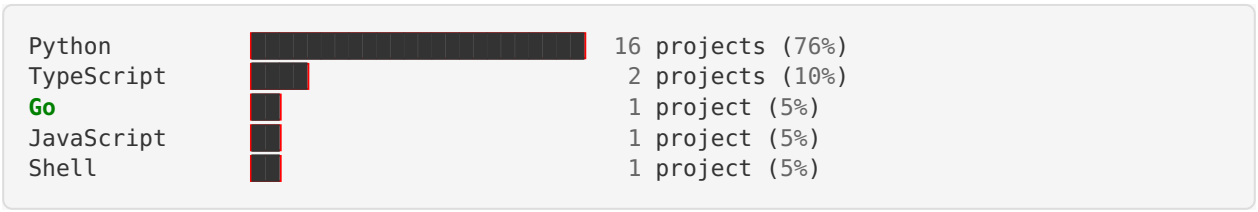
Project	Purpose	Key Capabilities
V1Databricks	Databricks Security	Three-layer security model, SIEM integration, Cyber Risk Index
AI-Intel-Feed	Threat Intelligence	AI software stack threat intel repository

📊 DATA & TRAINING DOMAIN (1 Project)

Project	Purpose	Key Capabilities
DatasetScrapper	Dataset Generation	Multi-role agents, instruct/ reasoning datasets, local AI

2.2 Common Technologies and Frameworks

Programming Languages Distribution



Core Frameworks & Technologies

Category	Technologies
AI/LLM	OpenAI API, Ollama, LM Studio, LiteLLM, Anthropic, Google AI
Web Frameworks	FastAPI, Next.js, Streamlit, Flask
Data Processing	Pandas, BeautifulSoup, Playwright, Crawl4ai
Security Standards	OWASP LLM Top 10, MITRE ATT&CK, MITRE ATLAS, OCSF
Detection Rules	SIGMA, YARA, Suricata
SIEM Integration	Trend Vision One, CEF format, Syslog
Containerization	Docker, Docker Compose, Kubernetes
Databases	PostgreSQL, SQLite, Redis
AI Frameworks	LangChain, CrewAI, LlamaIndex, Semantic Kernel

Shared Integration Points

1. **Trend Vision One** - 8 projects integrate with Vision One platform
2. **OWASP Standards** - 6 projects reference OWASP LLM Top 10
3. **Local LLM Support** - 12 projects support Ollama/LM Studio
4. **Docker Deployment** - 14 projects have containerization support

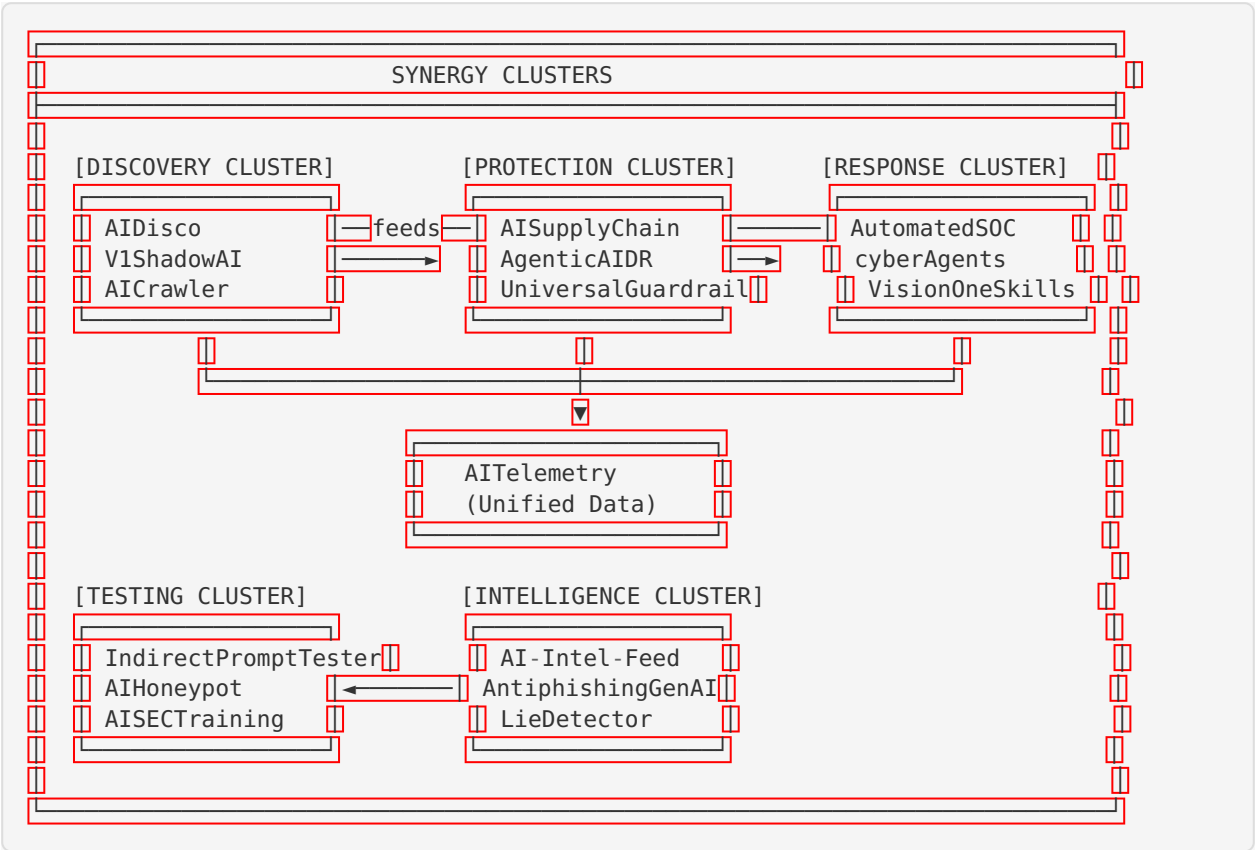
---

2.3 Overlapping Capabilities and Synergies

Capability Matrix

Capability	Projects Providing It
Shadow AI Detection	AIDisco, V1ShadowAI, AICrawler, VisionOneSkills
Prompt Injection Detection	IndirectPromptTester, AIHoneypot, AISECTraining, AntiphishingGenAI
Multi-Agent Architecture	AutomatedSOC, cyberAgents, AICrawler, DatasetScrapper
SIEM/XDR Integration	V1Databricks, AIDataGuard, AITelemetry, VisionOneSkills
Threat Intelligence	AI-Intel-Feed, cyberAgents, AICrawler, AntiphishingGenAI
Supply Chain Security	AISupplyChain, VisionOneSkills, AISECTraining
Agentic AI Security	AgenticAIDR, AutomatedSOC, cyberAgents

Natural Synergies



2.4 Integration Opportunities

Integration Pair	Benefit
AIDisco → AISupplyChain	Discovered AI assets feed into supply chain inventory
AICrawler → AI-Intel-Feed	Crawled AI services populate threat intel
AITelemetry → AutomatedSOC	Telemetry data enables intelligent triage
IndirectPromptTester → AgenticAIDR	Test results inform agent policies
VisionOneSkills → V1Databricks	Endpoint skills complement Databricks security
AntiphishingGenAI → LieDetector	Email analysis feeds social engineering detection
AIHoneypot → DatasetScrapper	Honeypot data generates training sets
cyberAgents → All Projects	Multi-agent analysis coordinates all components

---

3. Unified Project Proposal

3.1 Platform Name

AEGIS-AI

Autonomous Enterprise Guardian for Intelligent Security

Tagline: “Comprehensive AI Security for the AI Era”

Alternative Names Considered:
- SentinelAI - AI Security Command Center
- CyberNexus AI - Unified AI Security Platform
- AIShield Enterprise - Complete AI Threat Defense

3.2 Platform Vision

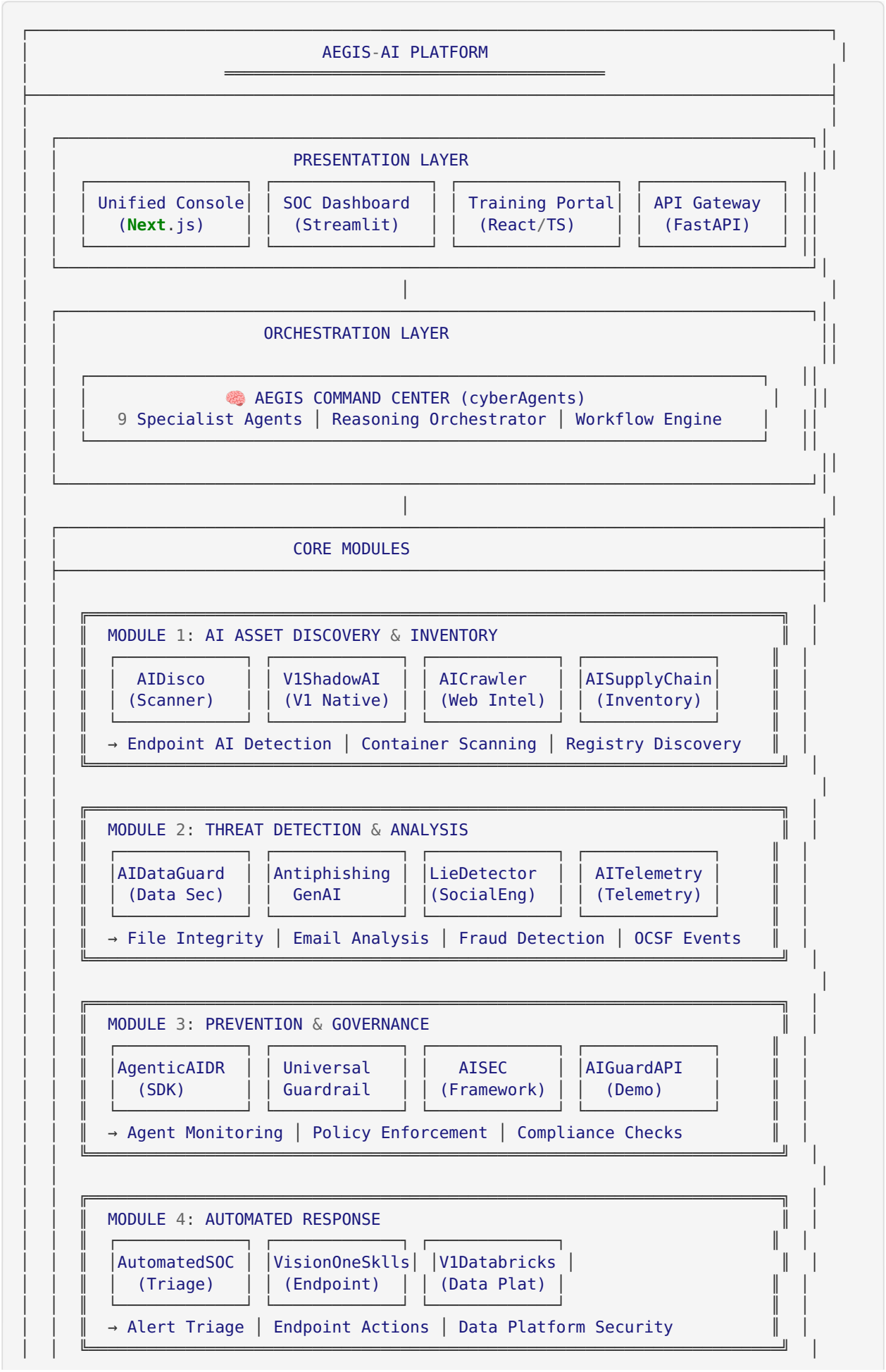
AEGIS-AI is a **production-ready, unified cybersecurity platform** that consolidates all 21 projects into a cohesive architecture, providing enterprises with:

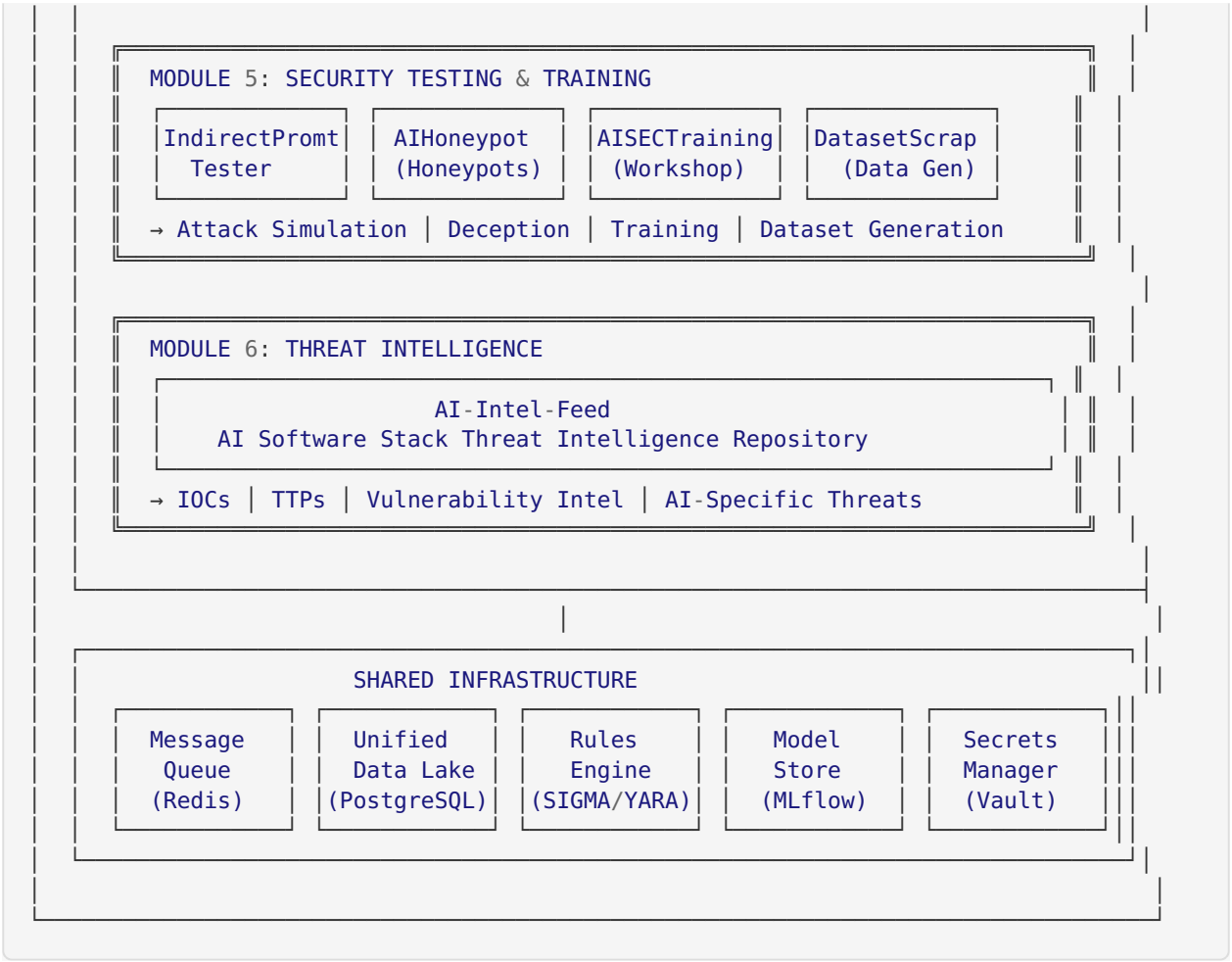
1. **Complete AI Asset Visibility** - Know every AI system in your environment
2. **Proactive AI Threat Prevention** - Stop attacks before they happen
3. **Intelligent Threat Detection** - AI-powered security operations

4. **Automated Response** - Agentic security that acts at machine speed
5. **Compliance & Governance** - Meet regulatory requirements for AI systems
6. **Continuous Security Validation** - Test and improve defenses

---

3.3 Modular Architecture - Project Mapping


AEGIS-AI PLATFORM
PRESENTATION LAYER
Unified Console (Next.js)
SOC Dashboard (Streamlit)
Training Portal (React/TS)
API Gateway (FastAPI)
ORCHESTRATION LAYER
AEGIS COMMAND CENTER (cyberAgents)
9 Specialist Agents | Reasoning Orchestrator | Workflow Engine
CORE MODULES
MODULE 1: AI ASSET DISCOVERY & INVENTORY
AIDisco (Scanner)
V1ShadowAI (V1 Native)
AICrawler (Web Intel)
AISupplyChain (Inventory)
→ Endpoint AI Detection | Container Scanning | Registry Discovery
MODULE 2: THREAT DETECTION & ANALYSIS
AIDataGuard (Data Sec)
Antiphishing GenAI
LieDetector (SocialEng)
AITelemetry (Telemetry)
→ File Integrity | Email Analysis | Fraud Detection | OCSF Events
MODULE 3: PREVENTION & GOVERNANCE
AgenticAIDR (SDK)
Universal Guardrail
AISEC (Framework)
AIGuardAPI (Demo)
→ Agent Monitoring | Policy Enforcement | Compliance Checks
MODULE 4: AUTOMATED RESPONSE
AutomatedSOC (Triage)
VisionOneSklls (Endpoint)
V1Databricks (Data Plat)
→ Alert Triage | Endpoint Actions | Data Platform Security


MODULE 5: SECURITY TESTING & TRAINING
IndirectPromt Tester
AIHoneypot (Honeypots)
AISECTraining (Workshop)
DatasetScrap (Data Gen)
→ Attack Simulation | Deception | Training | Dataset Generation
MODULE 6: THREAT INTELLIGENCE
AI-Intel-Feed
AI Software Stack Threat Intelligence Repository
→ IOCs | TTPs | Vulnerability Intel | AI-Specific Threats
SHARED INFRASTRUCTURE
Message Queue (Redis)
Unified Data Lake (PostgreSQL)
Rules Engine (SIGMA/YARA)
Model Store (MLflow)
Secrets Manager (Vault)

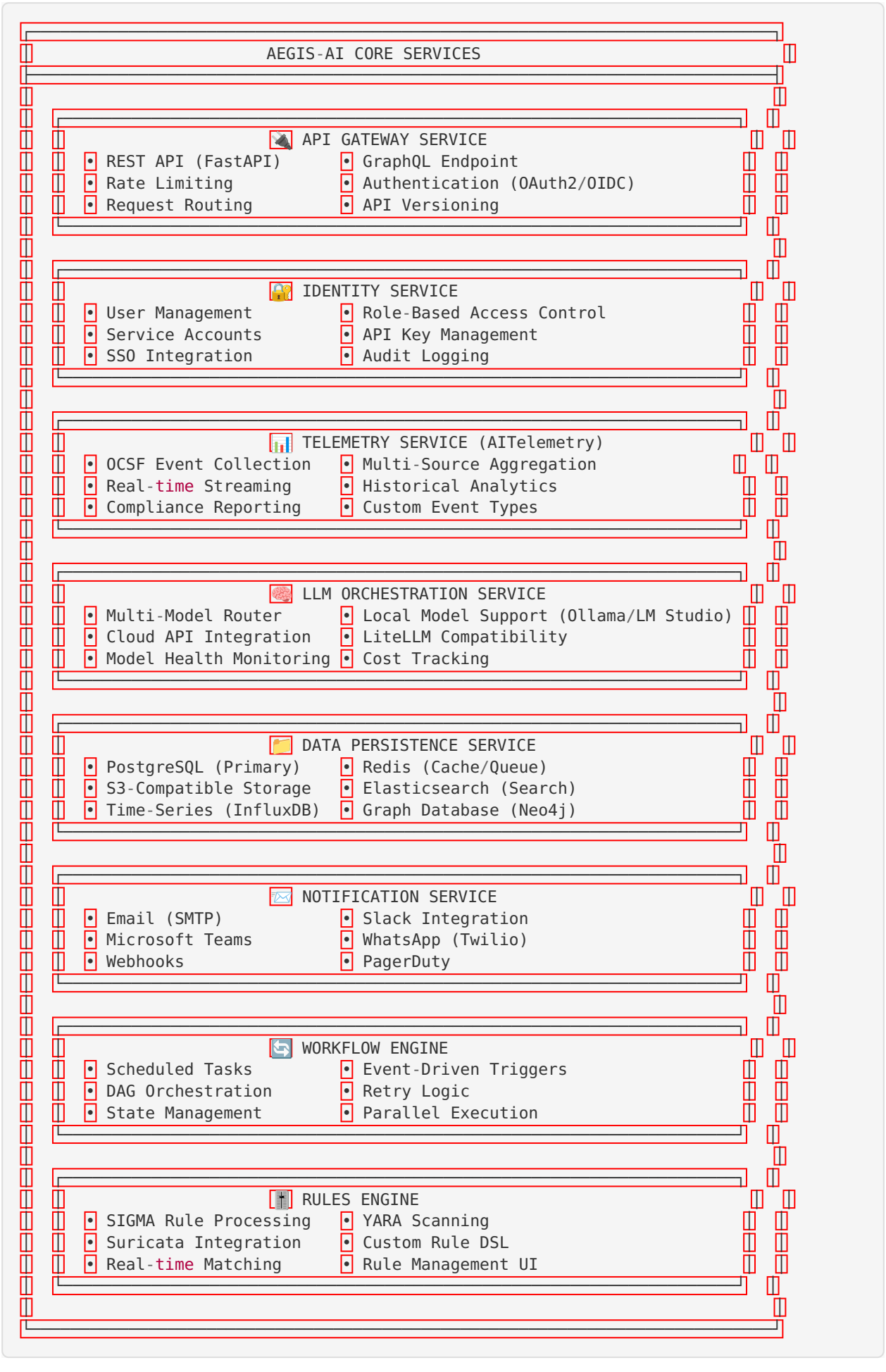
3.4 Feature Preservation Matrix

Original Project	Features Preserved in AEGIS-AI
V1Databricks	✓ Three-layer security model ✓ Audit log forwarding ✓ Cyber Risk Index ✓ AI app security monitoring
AISECTraining	✓ OWASP LLM Top 10 exercises ✓ Dual security approaches ✓ Progressive difficulty ✓ Multiple LLM support
AutomatedSOC	✓ Agentic triage ✓ Risk prioritization ✓ Response actions ✓ Multi-channel notifications ✓ Roster management
VisionOneSkills	✓ 7 production skills ✓ Cross-platform support ✓ MITRE mapping ✓ CEF output
AIDisco	✓ LLM software detection ✓ Docker/WSL2 scanning ✓ SIGMA rules ✓ Go binary deployment
AIDataGuard	✓ File integrity monitoring ✓ Log inspection rules ✓ Alert management ✓ Encrypted credentials
AICrawler	✓ Multi-agent discovery ✓ Risk scoring ✓ Detection rule generation ✓ Multiple AI backends
AI-Intel-Feed	✓ AI stack threat intel ✓ IOC repository
V1ShadowAI	✓ V1 native detection ✓ Existing control leverage
AITelemetry	✓ OCSF schema ✓ RFC standard ✓ Platform collectors ✓ Syslog/S3 forwarding
AISupplyChain	✓ AI-BOM ✓ Artifact signing ✓ Lineage tracking ✓ CoSAI controls ✓ Vendor risk
IndirectPromptTester	✓ 118 attack examples ✓ 21 vectors ✓ File generation ✓ Difficulty grading
AntiphishingGenAI	✓ Linguistic analysis ✓ AI content detection ✓ Attachment scanning ✓ Email integrations
AgenticAIDR	✓ Agent instrumentation ✓ Real-time tracking ✓ Policy enforcement ✓ Kill-switch ✓ 9+ framework support

Original Project	Features Preserved in AEGIS-AI
AIHoneypot	✓ OWASP LLM demos ✓ Honeypot deployment ✓ Security training
cyberAgents	✓ 9 specialist agents ✓ PHI-4 orchestrator ✓ Local model support ✓ TI integrations
AISEC	✓ Security framework ✓ Best practices
AIGuardAPIDemo	✓ Guard API examples
LieDetector	✓ Trust scoring ✓ Social engineering detection ✓ Behavioral nudges
UniversalGuardrail	✓ Guardrail standard ✓ Universal API
DatasetScrapper	✓ Multi-role agents ✓ Instruct datasets ✓ Reasoning datasets ✓ Local AI

3.5 Core Platform Services

Service Layer Architecture


AEGIS-AI CORE SERVICES
API GATEWAY SERVICE
REST API (FastAPI)
GraphQL Endpoint
Rate Limiting
Authentication (OAuth2/OIDC)
Request Routing
API Versioning
IDENTITY SERVICE
User Management
Role-Based Access Control
Service Accounts
API Key Management
SSO Integration
Audit Logging
TELEMETRY SERVICE (AITelemetry)
OCSF Event Collection
Multi-Source Aggregation
Real-time Streaming
Historical Analytics
Compliance Reporting
Custom Event Types
LLM ORCHESTRATION SERVICE
Multi-Model Router
Local Model Support (Ollama/LM Studio)
Cloud API Integration
LiteLLM Compatibility
Model Health Monitoring
Cost Tracking
DATA PERSISTENCE SERVICE
PostgreSQL (Primary)
Redis (Cache/Queue)
S3-Compatible Storage
Elasticsearch (Search)
Time-Series (InfluxDB)
Graph Database (Neo4j)
NOTIFICATION SERVICE
Email (SMTP)
Slack Integration
Microsoft Teams
WhatsApp (Twilio)
Webhooks
PagerDuty
WORKFLOW ENGINE
Scheduled Tasks
Event-Driven Triggers
DAG Orchestration
Retry Logic
State Management
Parallel Execution
RULES ENGINE
SIGMA Rule Processing
YARA Scanning
Suricata Integration
Custom Rule DSL
Real-time Matching
Rule Management UI

3.6 Deployment Considerations

Deployment Modes

Mode	Description	Use Case
Cloud-Native	Full Kubernetes deployment on AWS/GCP/Azure	Enterprise production
Hybrid	Core in cloud, agents on-premise	Regulated industries
On-Premise	Complete self-hosted deployment	Air-gapped environments
Appliance	Pre-configured VM/container bundle	SMB quick-start
SaaS	Multi-tenant managed service	MSP/MSSP offering

Container Architecture

```yaml
# docker-compose.yml (simplified)
services:
  # Core Services
  api-gateway:
    image: aegis-ai/gateway:latest
    ports: ["8080:8080"]

  orchestrator:
    image: aegis-ai/orchestrator:latest  # cyberAgents
    depends_on: [redis, postgres]

  telemetry-collector:
    image: aegis-ai/telemetry:latest  # AITelemetry

  # Module Services
  asset-discovery:
    image: aegis-ai/discovery:latest  # AIDisco + AICrawler

  threat-detection:
    image: aegis-ai/detection:latest  # AIDataGuard + Antiphishing

  soc-automation:
    image: aegis-ai/soc:latest  # AutomatedSOC

  agentic-dr:
    image: aegis-ai/agentic-dr:latest  # AgenticAIDR

  supply-chain:
    image: aegis-ai/supply-chain:latest  # AISupplyChain

  security-testing:
    image: aegis-ai/testing:latest  # IndirectPromptTester + AIHoneypot

  training-portal:
    image: aegis-ai/training:latest  # AISECTraining

  # Infrastructure
  postgres:
    image: postgres:16-alpine

  redis:
    image: redis:7-alpine

  ollama:
    image: ollama/ollama:latest  # Local LLM
```

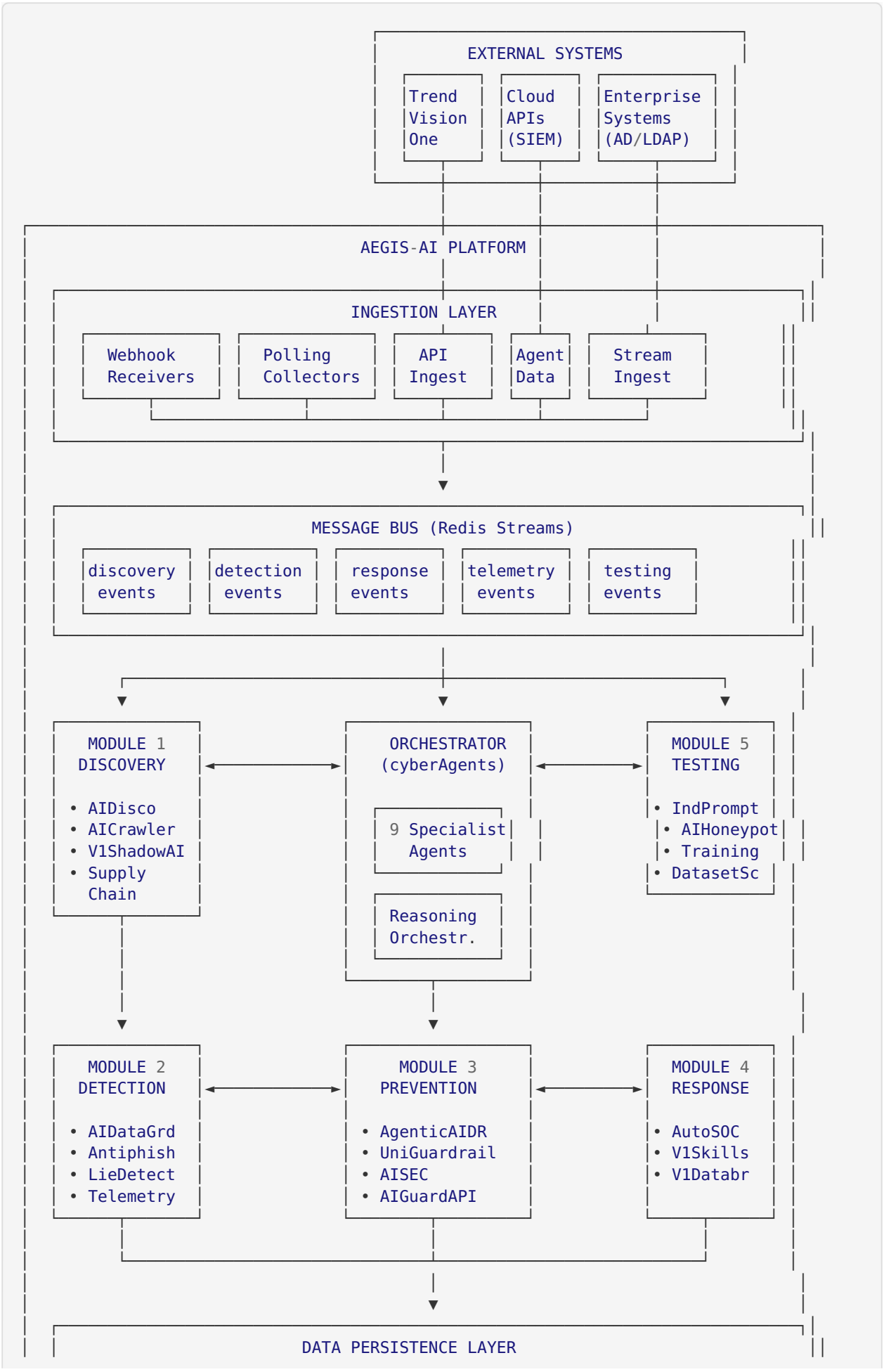
Operational Considerations

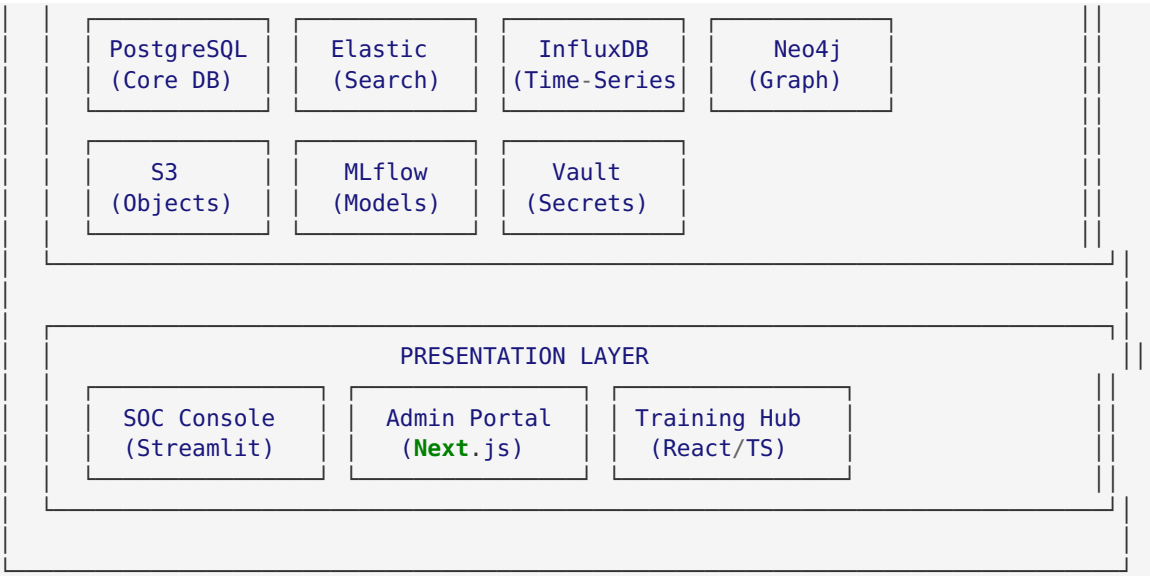
Aspect	Recommendation
Scaling	Horizontal pod autoscaling for stateless services
High Availability	Multi-zone deployment, database replication
Monitoring	Prometheus + Grafana stack, custom AI metrics
Logging	Centralized logging with ELK/Loki
Backup	Automated database backups, config versioning
Updates	Blue-green deployments, canary releases
Security	Network policies, secrets management (Vault)

---

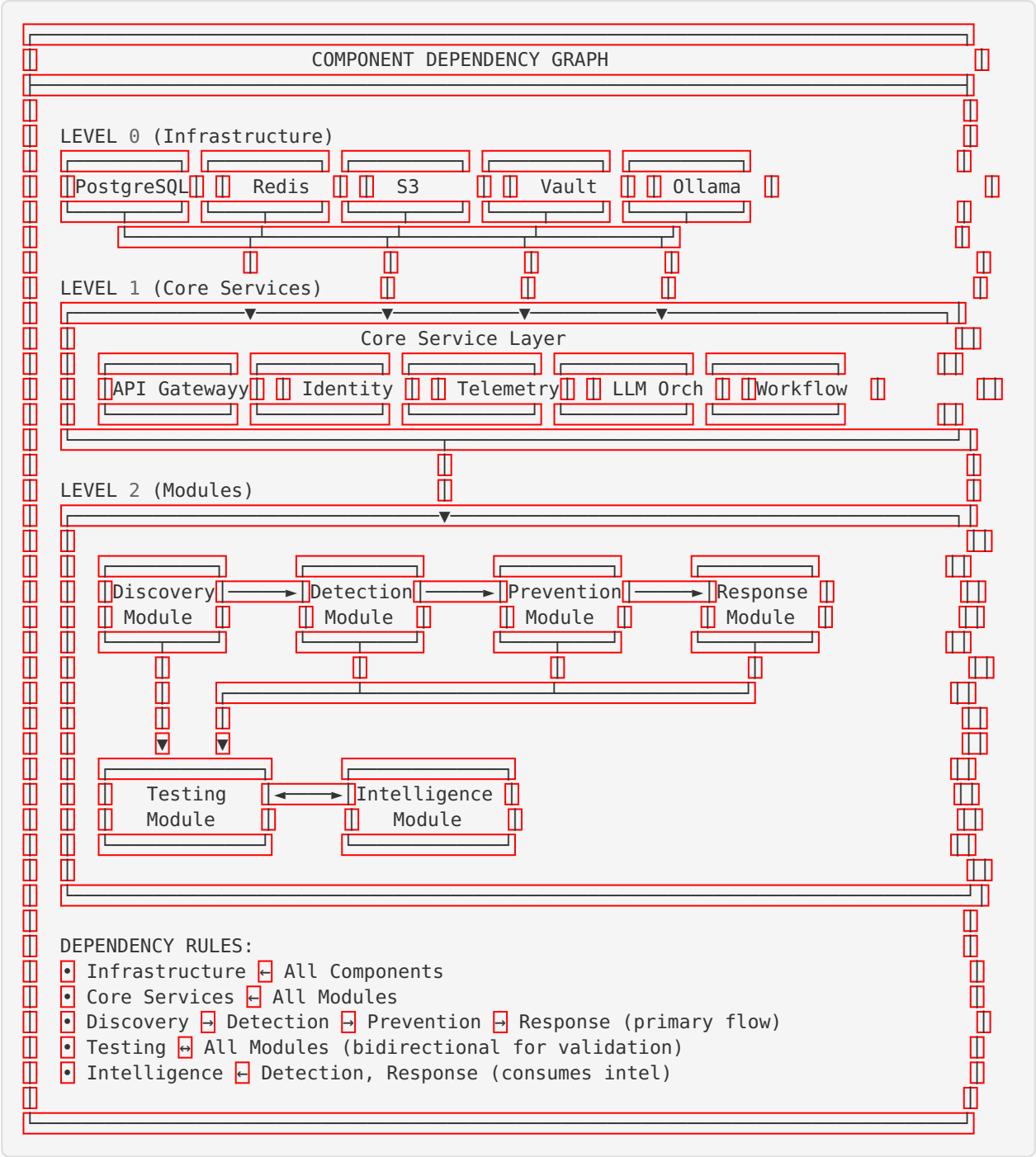
4. Technical Architecture

4.1 High-Level System Architecture

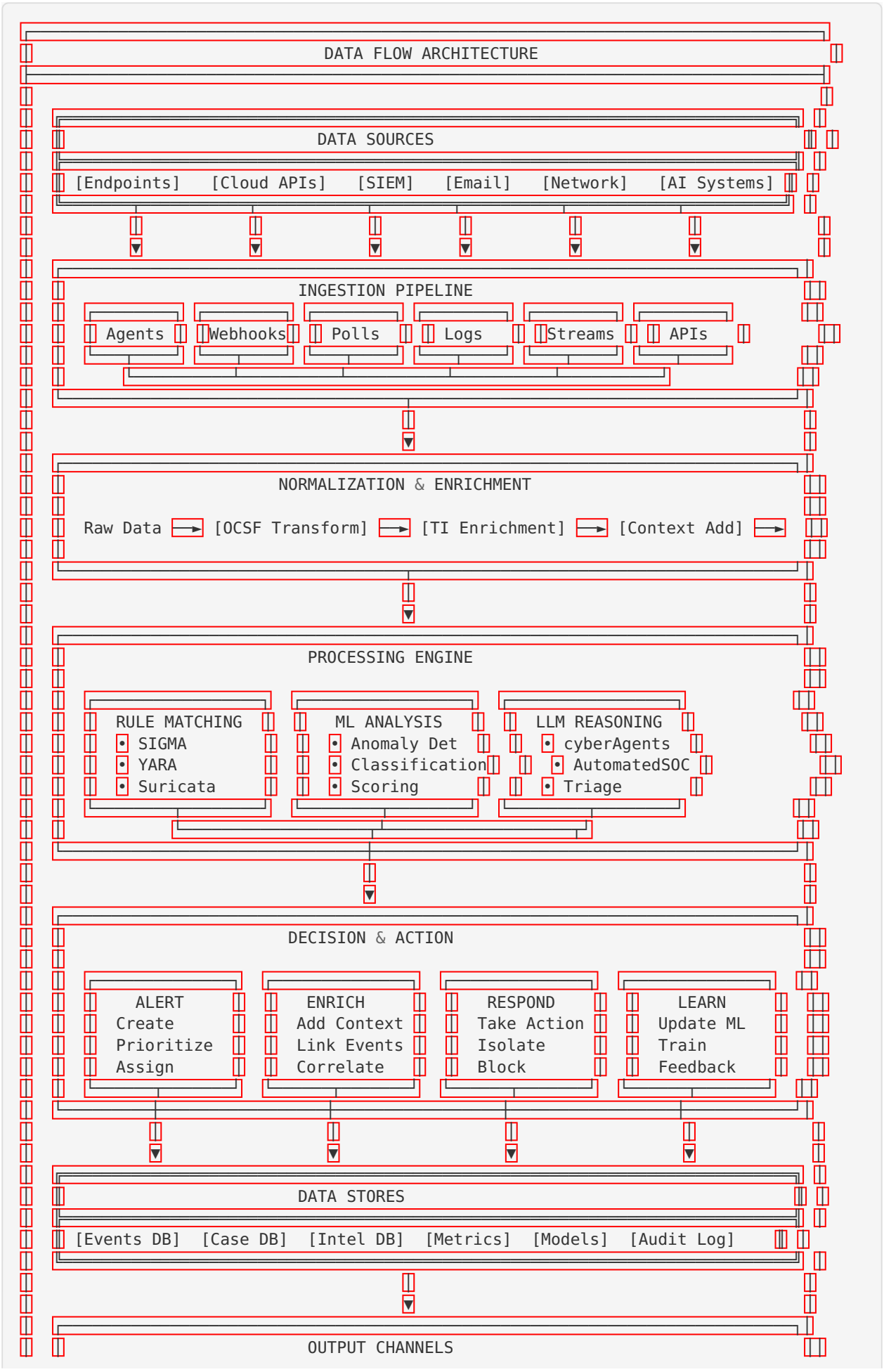

EXTERNAL SYSTEMS
Trend Vision One
Cloud APIs (SIEM)
Enterprise Systems (AD/LDAP)
AEGIS-AI PLATFORM
INGESTION LAYER
Webhook Receivers
Polling Collectors
API Ingest
Agent Data
Stream Ingest
MESSAGE BUS (Redis Streams)
discovery events
detection events
response events
telemetry events
testing events
MODULE 1 DISCOVERY
• AIDisco
• AICrawler
• V1ShadowAI
• Supply Chain
ORCHESTRATOR (cyberAgents)
9 Specialist Agents
Reasoning Orchestr.
MODULE 5 TESTING
• IndPrompt
• AIHoneypot
• Training
• DatasetSc
MODULE 2 DETECTION
• AIDataGrd
• Antiphish
• LieDetect
• Telemetry
MODULE 3 PREVENTION
• AgenticAIDR
• UniGuardrail
• AISEC
• AIGuardAPI
MODULE 4 RESPONSE
• AutoSOC
• V1Skills
• V1Databr
DATA PERSISTENCE LAYER


PostgreSQL
(Core DB)
Elastic
(Search)
InfluxDB
(Time-Series
Neo4j
(Graph)
S3
(Objects)
MLflow
(Models)
Vault
(Secrets)
PRESENTATION LAYER
SOC Console
(Streamlit)
Admin Portal
(Next.js)
Training Hub
(React/TS)

4.2 Component Relationships and Dependencies



4.3 Data Flow Diagram


DATA FLOW ARCHITECTURE
DATA SOURCES
[Endpoints]
[Cloud APIs]
[SIEM]
[Email]
[Network]
[AI Systems]
INGESTION PIPELINE
Agents
Webhooks
Polls
Logs
Streams
APIs
NORMALIZATION & ENRICHMENT
Raw Data
[OCSF Transform]
[TI Enrichment]
[Context Add]
PROCESSING ENGINE
RULE MATCHING
• SIGMA
• YARA
• Suricata
ML ANALYSIS
• Anomaly Det
• Classification
• Scoring
LLM REASONING
• cyberAgents
• AutomatedSOC
• Triage
DECISION & ACTION
ALERT
Create
Prioritize
Assign
ENRICH
Add Context
Link Events
Correlate
RESPOND
Take Action
Isolate
Block
LEARN
Update ML
Train
Feedback
DATA STORES
[Events DB]
[Case DB]
[Intel DB]
[Metrics]
[Models]
[Audit Log]
OUTPUT CHANNELS


UI
API
SIEM
Email
Slack
Webhooks

4.4 API/Integration Layer Design

REST API Structure

```
AEGIS-AI API v1
===============

Base URL: https://api.aegis-ai.io/v1

Authentication: Bearer Token (JWT) or API Key

┌─────────────────────────────────────────────────────────────────────────────┐
│                              API ENDPOINTS                                  │
├─────────────────────────────────────────────────────────────────────────────┤
│                                                                             │
│  DISCOVERY MODULE (/discovery)                                              │
│  ├── POST   /scan              # Trigger AI asset discovery scan            │
│  ├── GET    /assets            # List discovered AI assets                  │
│  ├── GET    /assets/{id}       # Get asset details                          │
│  ├── POST   /assets/search     # Search assets with filters                 │
│  └── GET    /inventory         # Get AI-BOM inventory                       │
│                                                                             │
│  DETECTION MODULE (/detection)                                              │
│  ├── POST   /analyze           # Analyze content/email/file                 │
│  ├── GET    /alerts            # List security alerts                       │
│  ├── GET    /alerts/{id}       # Get alert details                          │
│  ├── PATCH  /alerts/{id}       # Update alert status                        │
│  ├── POST   /rules             # Create detection rule                      │
│  └── GET    /telemetry         # Query telemetry events                     │
│                                                                             │
│  PREVENTION MODULE (/prevention)                                            │
│  ├── POST   /policies          # Create security policy                     │
│  ├── GET    /policies          # List policies                              │
│  ├── POST   /evaluate          # Evaluate agent action against policy       │
│  ├── POST   /guardrail/check   # Check content against guardrails           │
│  └── GET    /compliance        # Get compliance status                      │
│                                                                             │
│  RESPONSE MODULE (/response)                                                │
│  ├── POST   /triage            # Submit alert for AI triage                 │
│  ├── POST   /actions/isolate   # Isolate endpoint                           │
│  ├── POST   /actions/block     # Block hash/IP/domain                       │
│  ├── POST   /playbooks         # Execute response playbook                  │
│  └── GET    /cases             # List investigation cases                   │
│                                                                             │
│  TESTING MODULE (/testing)                                                  │
│  ├── POST   /prompts/generate  # Generate test prompt files                │
│  ├── GET    /prompts/database  # Query prompt injection database           │
│  ├── POST   /honeypot/deploy   # Deploy AI honeypot                         │
│  └── POST   /validate          # Validate security controls                 │
│                                                                             │
│  INTELLIGENCE MODULE (/intel)                                               │
│  ├── GET    /threats           # Get AI threat intelligence                 │
│  ├── POST   /iocs              # Submit IOC for analysis                    │
│  ├── GET    /vulnerabilities   # List AI vulnerabilities                    │
│  └── POST   /research          # Submit new threat research                 │
│                                                                             │
│  AGENTS MODULE (/agents)                                                    │
│  ├── POST   /query             # Query cyberAgents for analysis             │
│  ├── GET    /agents            # List available specialist agents           │
│  ├── POST   /workflow          # Execute multi-agent workflow               │
│  └── GET    /sessions/{id}     # Get agent session details                  │
│                                                                             │
│  TRAINING MODULE (/training)                                                │
│  ├── GET    /exercises         # List training exercises                    │
│  ├── POST   /exercises/start   # Start training exercise                    │
│  ├── GET    /progress          # Get user training progress                 │
```

```
│  └── POST   /datasets/generate # Generate training datasets                  │
│                                                                              │
│  SYSTEM (/system)                                                            │
│  ├── GET    /health            # Health check                                │
│  ├── GET    /metrics           # Prometheus metrics                          │
│  ├── GET    /status            # System status dashboard                     │
│  └── GET    /config            # Get system configuration                    │
│                                                                              │
└──────────────────────────────────────────────────────────────────────────────┘
```

Event Schema (OCSF-Compatible)

```json
{
  "metadata": {
    "version": "1.0.0",
    "product": "AEGIS-AI",
    "profiles": ["security"],
    "uid": "evt_abc123xyz",
    "logged_time": "2026-02-12T15:30:00Z"
  },
  "class_uid": 2001,
  "class_name": "Security Finding",
  "category_uid": 2,
  "category_name": "Findings",
  "severity_id": 4,
  "severity": "High",
  "type_uid": 200101,
  "type_name": "AI Security Finding",
  "activity_id": 1,
  "activity_name": "Detection",
  "message": "Shadow AI installation detected",
  "finding": {
    "title": "Unauthorized Ollama Installation",
    "desc": "Local LLM server detected on endpoint",
    "types": ["shadow_ai", "policy_violation"],
    "related_events": [],
    "remediation": {
      "desc": "Review and approve or remove installation",
      "kb_articles": ["KB-AI-001"]
    },
    "src_product": "AIDisco"
  },
  "device": {
    "hostname": "workstation-001",
    "ip": "10.0.1.50",
    "os": {"name": "Windows 11", "type_id": 100}
  },
  "risk_score": 75,
  "enrichments": [
    {"name": "asset_criticality", "value": "high"},
    {"name": "data_sensitivity", "value": "pii_present"}
  ]
}
```

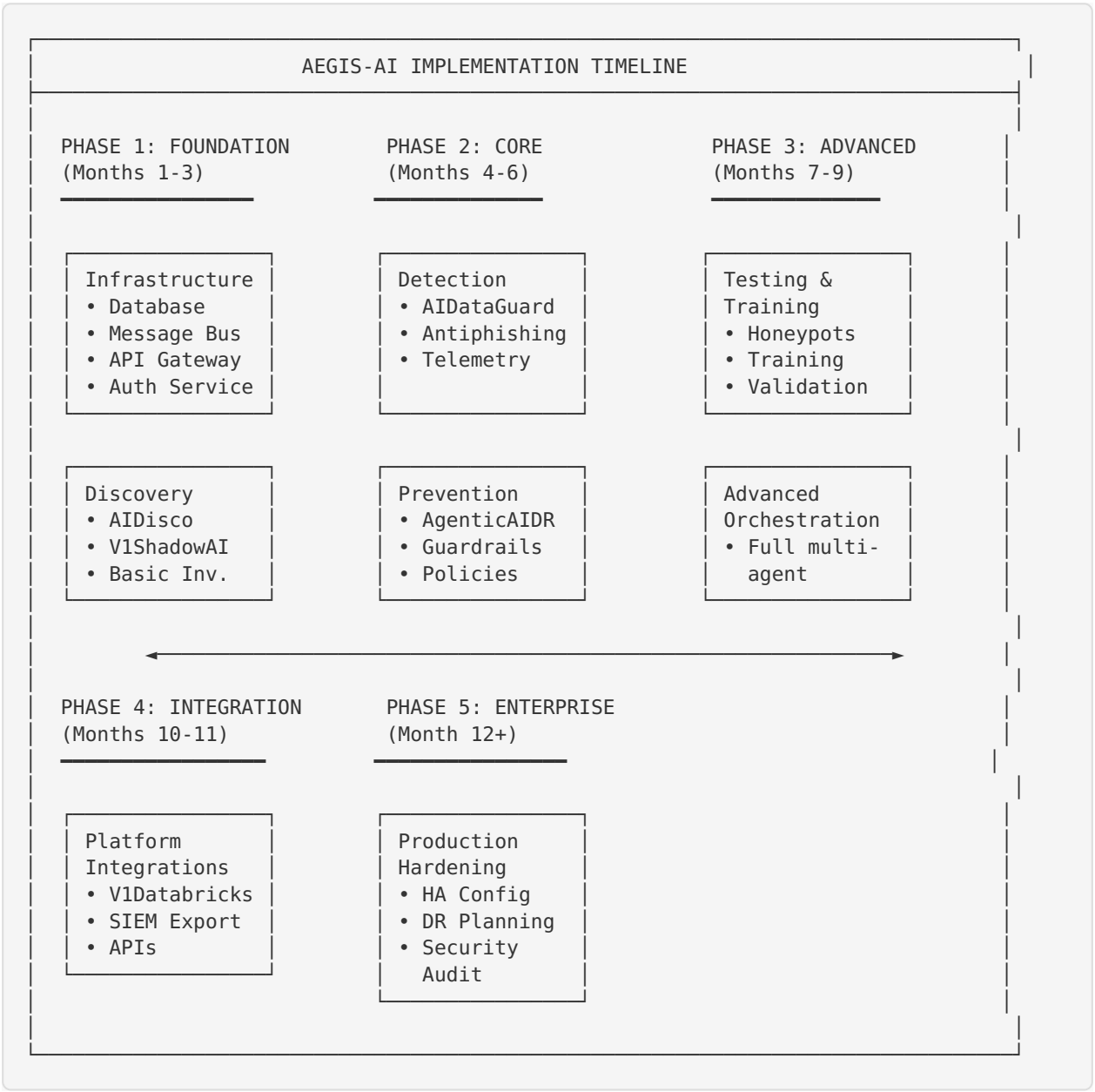
4.5 Technology Stack Recommendations

Layer	Technology	Rationale
API Gateway	FastAPI + Uvicorn	Async Python, automatic OpenAPI
Frontend - SOC	Streamlit	Rapid development, data-rich
Frontend - Admin	Next.js 14	Production-grade React
Frontend - Training	React + TypeScript	Interactive exercises
Primary Database	PostgreSQL 16	ACID, JSON support, mature
Cache/Queue	Redis 7	Streams for events, caching
Search Engine	Elasticsearch 8	Log search, analytics
Time Series	InfluxDB 3	Metrics, telemetry
Graph Database	Neo4j	Asset relationships
Object Storage	MinIO (S3-compatible)	Models, artifacts
ML Platform	MLflow	Model registry, tracking
LLM Orchestration	LiteLLM	Multi-provider routing
Local LLM	Ollama	Air-gapped deployments
Container Orchestration	Kubernetes (K8s)	Production scaling
CI/CD	GitHub Actions	Automated pipelines
Monitoring	Prometheus + Grafana	Metrics, dashboards
Secrets	HashiCorp Vault	Secure credentials

---

5. Implementation Roadmap

5.1 Phased Approach Overview



5.2 Detailed Phase Breakdown

Phase 1: Foundation (Months 1-3)

Sprint	Focus	Deliverables	Projects Integrated
1.1	Infrastructure Setup	K8s cluster, databases, message bus	-
1.2	Core Services	API Gateway, Identity, base APIs	-
1.3	Discovery Module	Basic AI asset scanning	AIDisco, V1ShadowAI
1.4	Data Layer	Schema design, migrations, initial ETL	AITelemetry (schema)
1.5	Basic UI	SOC dashboard skeleton, admin portal	-
1.6	CI/CD Pipeline	Automated testing, deployment	-

Milestone: MVP with basic discovery and visibility

Phase 2: Core Detection & Prevention (Months 4-6)

Sprint	Focus	Deliverables	Projects Integrated
2.1	Detection Engine	Rule processing, SIGMA support	AIDataGuard
2.2	Email Security	GenAI phishing detection	AntiphishingGenAI
2.3	Social Engineering	Trust scoring, behavioral analysis	LieDetector
2.4	Telemetry Pipeline	Full OCSF collection, forwarding	AITelemetry
2.5	Agentic Security	Agent instrumentation SDK	AgenticAIDR
2.6	Policy Engine	Guardrails, enforcement	UniversalGuardrail, AISEC

Milestone: Complete detection and prevention capabilities

Phase 3: Advanced Capabilities (Months 7-9)

Sprint	Focus	Deliverables	Projects Integrated
3.1	Security Testing	Prompt injection framework	IndirectPromptTester
3.2	Honeypot System	Deployable vulnerable apps	AIHoneypot
3.3	Training Platform	Full workshop integration	AISECTraining
3.4	Dataset Generation	Training data creation	DatasetScrapper
3.5	Supply Chain	AI-BOM, signing, provenance	AISupplyChain
3.6	Web Intelligence	AI service crawler	AICrawler

Milestone: Complete testing, training, and supply chain modules

---

Phase 4: Integration & Automation (Months 10-11)

Sprint	Focus	Deliverables	Projects Integrated
4.1	SOC Automation	Full agentic triage	AutomatedSOC
4.2	Endpoint Skills	Cross-platform skill deployment	VisionOneSkills
4.3	Platform Integration	Databricks security layer	V1Databricks
4.4	Multi-Agent	Full cyberAgents integration	cyberAgents
4.5	Intelligence Feed	AI threat intel repository	AI-Intel-Feed
4.6	Unified Dashboard	Complete SOC experience	-

Milestone: Full automation and integration capabilities

Phase 5: Enterprise Hardening (Month 12+)

Focus Area	Activities
High Availability	Multi-zone deployment, database replication, failover testing
Disaster Recovery	Backup strategies, recovery procedures, DR drills
Performance	Load testing, optimization, caching strategies
Security Audit	Penetration testing, code review, compliance certification
Documentation	Admin guides, API docs, runbooks, training materials
Support Model	Incident procedures, escalation paths, SLAs

Milestone: Production-ready enterprise platform

5.3 Resource Requirements

Role	Count	Phase
Platform Architect	1	All
Backend Engineers (Python)	3-4	All
Frontend Engineers	2	1-4
DevOps/SRE	2	All
ML/AI Engineer	1	2-4
Security Engineer	1	3-5
QA Engineer	1	2-5
Technical Writer	1	4-5

5.4 Risk Mitigation

Risk	Mitigation Strategy
Integration complexity	Modular design with clear interfaces, extensive testing
LLM reliability	Multi-model fallback, local model support
Performance at scale	Async processing, horizontal scaling, caching
Security vulnerabilities	Regular audits, secure development practices
Vendor lock-in	Abstraction layers, open standards (OCSF, SIGMA)
Scope creep	Strict phase gates, MVP-first approach

Appendix A: Project Quick Reference

#	Project	Primary Function	Primary Language	V1 Integration
1	V1Databricks	Databricks Security	Python	✅
2	AISECTraining	Security Training	TypeScript	✅
3	AutomatedSOC	SOC Triage	Python	✅
4	VisionOneSkills	Endpoint Skills	Python/Shell	✅
5	AIDisco	AI Scanner	Python/Go	❌
6	AIDataGuard	Data Security	TypeScript	✅
7	AICrawler	AI Discovery	Python	❌
8	AI-Intel-Feed	Threat Intel	-	❌
9	V1ShadowAI	Shadow AI Detection	Python	✅
10	AITelemetry	Telemetry Standard	Python	❌
11	AISupplyChain	Supply Chain	Python	❌
12	IndirectPrompt-Tester	Prompt Testing	Python	❌
13	Antiphishing-GenAI	Phishing Detection	Python	❌
14	AgenticAIDR	Agentic AI D&R	Python	❌
15	AIHoneypot	Honeypots	Python	❌
16	cyberAgents	Multi-Agent	Python	✅
17	AISEC	Security Framework	Python	❌
18	AIGuardAPIDemo	API Demo	-	✅
19	LieDetector	Social Engineering	Python/JS	❌
20			-	❌

#	Project	Primary Function	Primary Language	V1 Integration
	UniversalGuardrail	Guardrail Standard		
21	DatasetScrapper	Dataset Generation	Python	❌

Appendix B: Glossary

Term	Definition
AEGIS-AI	Proposed unified platform name (Autonomous Enterprise Guardian for Intelligent Security)
AI-BOM	AI Bill of Materials - inventory of AI components
CoSAI	Coalition for Secure AI
MITRE ATLAS	Adversarial Threat Landscape for AI Systems
OCSF	Open Cybersecurity Schema Framework
Shadow AI	Unauthorized/unmanaged AI deployments
Vision One	Trend Micro’s unified security platform

Document generated by DeepAgent - February 12, 2026