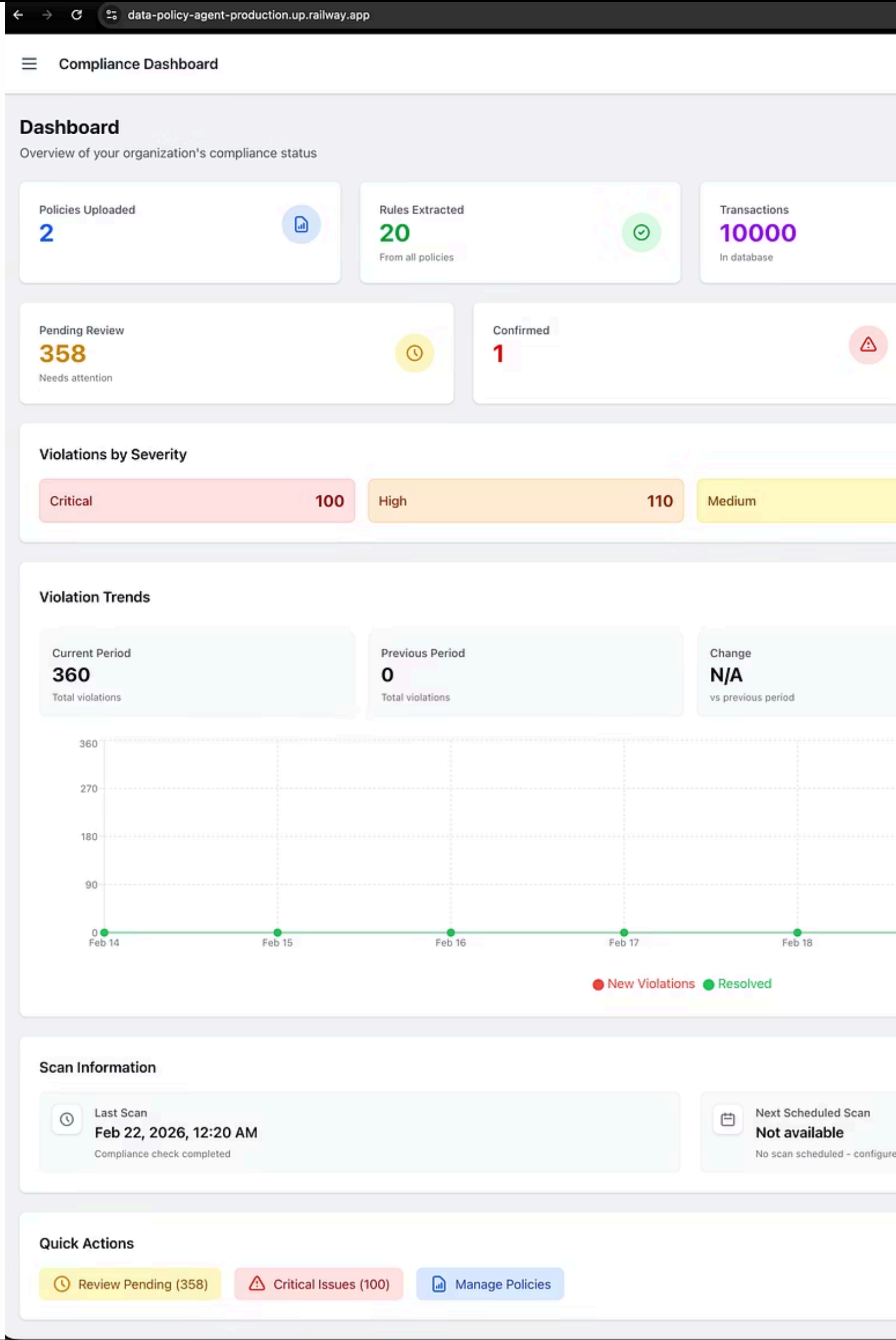
Data Policy Agent

From Static Compliance PDFs to Real-Time Enforcement

Autonomous Compliance Intelligence – transforming regulatory documents into executable monitoring systems that detect violations in real-time.

OPEN SOURCE AI-POWERED



data-policy-agent-production.up.railway.app

Compliance Dashboard

Dashboard

Overview of your organization's compliance status

Policies Uploaded
2

Rules Extracted
20
From all policies

Transactions
10000
In database

Pending Review
358
Needs attention

Confirmed
1

Violations by Severity

Critical 100

High 110

Medium

Violation Trends

Current Period
360
Total violations

Previous Period
0
Total violations

Change
N/A
vs previous period

Scan Information

Last Scan
Feb 22, 2026, 12:20 AM
Compliance check completed

Next Scheduled Scan
Not available
No scan scheduled - configure

Quick Actions

Review Pending (358) Critical Issues (100) Manage Policies

The Compliance Crisis

Organizations face a fundamental disconnect between their compliance obligations and enforcement capabilities. Regulatory policies exist as static PDF documents that humans must interpret and manually enforce through spreadsheet-based audits.

Document Paralysis

Compliance policies stored as unstructured PDFs and Word documents, inaccessible to automated systems and requiring constant manual interpretation.

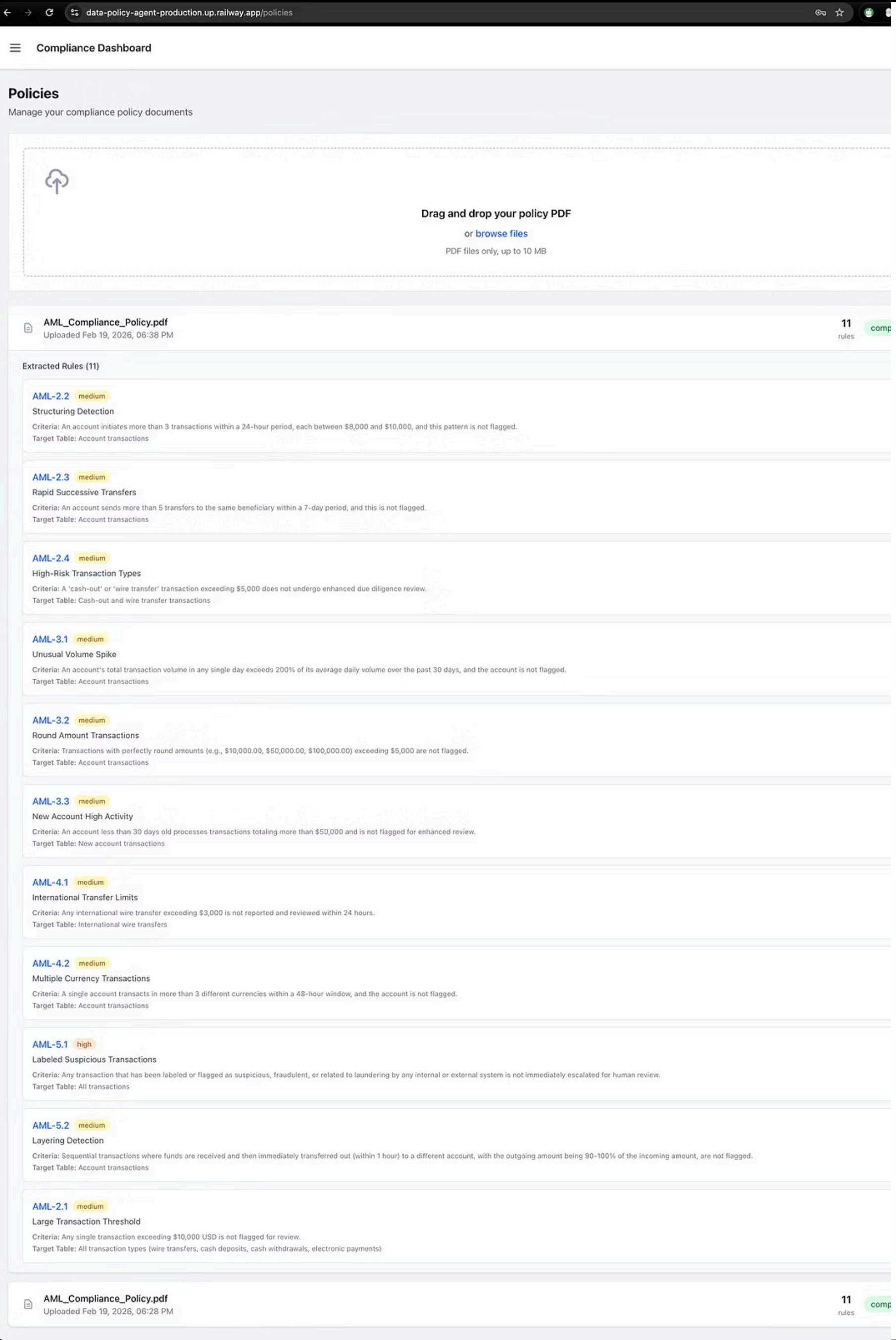
Audit Lag

Manual spreadsheet reviews conducted weekly or monthly, creating dangerous gaps where violations accumulate undetected in high-volume environments.

Regulatory Exposure

Delayed violation detection leads to financial penalties, regulatory sanctions, and reputational damage across AML, fraud prevention, and internal compliance.

The cost of non-compliance extends beyond fines—organizations risk operational shutdowns, loss of licenses, and erosion of stakeholder trust. Traditional approaches cannot scale with modern transaction volumes.



Policies

Manage your compliance policy documents

Drag and drop your policy PDF

or browse files

PDF files only, up to 10 MB

AML_Compliance_Policy.pdf
Uploaded Feb 19, 2026, 06:38 PM

11 rules comp

Extracted Rules (11)

AML-2.2 medium
Structuring Detection
Criteria: An account initiates more than 3 transactions within a 24-hour period, each between $8,000 and $10,000, and this pattern is not flagged.
Target Table: Account transactions

AML-2.3 medium
Rapid Successive Transfers
Criteria: An account sends more than 5 transfers to the same beneficiary within a 7-day period, and this is not flagged.
Target Table: Account transactions

AML-2.4 medium
High-Risk Transaction Types
Criteria: A 'cash-out' or 'wire transfer' transaction exceeding $5,000 does not undergo enhanced due diligence review.
Target Table: Cash-out and wire transfer transactions

AML-3.1 medium
Unusual Volume Spike
Criteria: An account's total transaction volume in any single day exceeds 200% of its average daily volume over the past 30 days, and the account is not flagged.
Target Table: Account transactions

AML-3.2 medium
Round Amount Transactions
Criteria: Transactions with perfectly round amounts (e.g., $10,000.00, $50,000.00, $100,000.00) exceeding $5,000 are not flagged.
Target Table: Account transactions

AML-3.3 medium
New Account High Activity
Criteria: An account less than 30 days old processes transactions totaling more than $50,000 and is not flagged for enhanced review.
Target Table: New account transactions

AML-4.1 medium
International Transfer Limits
Criteria: Any international wire transfer exceeding $3,000 is not reported and reviewed within 24 hours.
Target Table: International wire transfers

AML-4.2 medium
Multiple Currency Transactions
Criteria: A single account transacts in more than 3 different currencies within a 48-hour window, and the account is not flagged.
Target Table: Account transactions

AML-5.1 high
Labeled Suspicious Transactions
Criteria: Any transaction that has been labeled or flagged as suspicious, fraudulent, or related to laundering by any internal or external system is not immediately escalated for human review.
Target Table: All transactions

AML-5.2 medium
Layering Detection
Criteria: Sequential transactions where funds are received and then immediately transferred out (within 1 hour) to a different account, with the outgoing amount being 90-100% of the incoming amount, are not flagged.
Target Table: Account transactions

AML-2.1 medium
Large Transaction Threshold
Criteria: Any single transaction exceeding $10,000 USD is not flagged for review.
Target Table: All transaction types (wire transfers, cash deposits, cash withdrawals, electronic payments)

AML_Compliance_Policy.pdf
Uploaded Feb 19, 2026, 06:28 PM

11 rules comp

Solution: Executable Compliance Logic

Data Policy Agent converts compliance documents into executable monitoring systems through an AI-powered pipeline that bridges the gap between regulatory text and real-time enforcement.

Upload Policy PDF
Secure document ingestion with version tracking

AI Rule Extraction
NLP converts unstructured text to structured logic

Rule Storage
Persistent rule repository with audit trail

Database Connection
Secure PostgreSQL integration

Transaction Scan
Automated rule evaluation at scale

Violation Detection
Real-time identification and classification

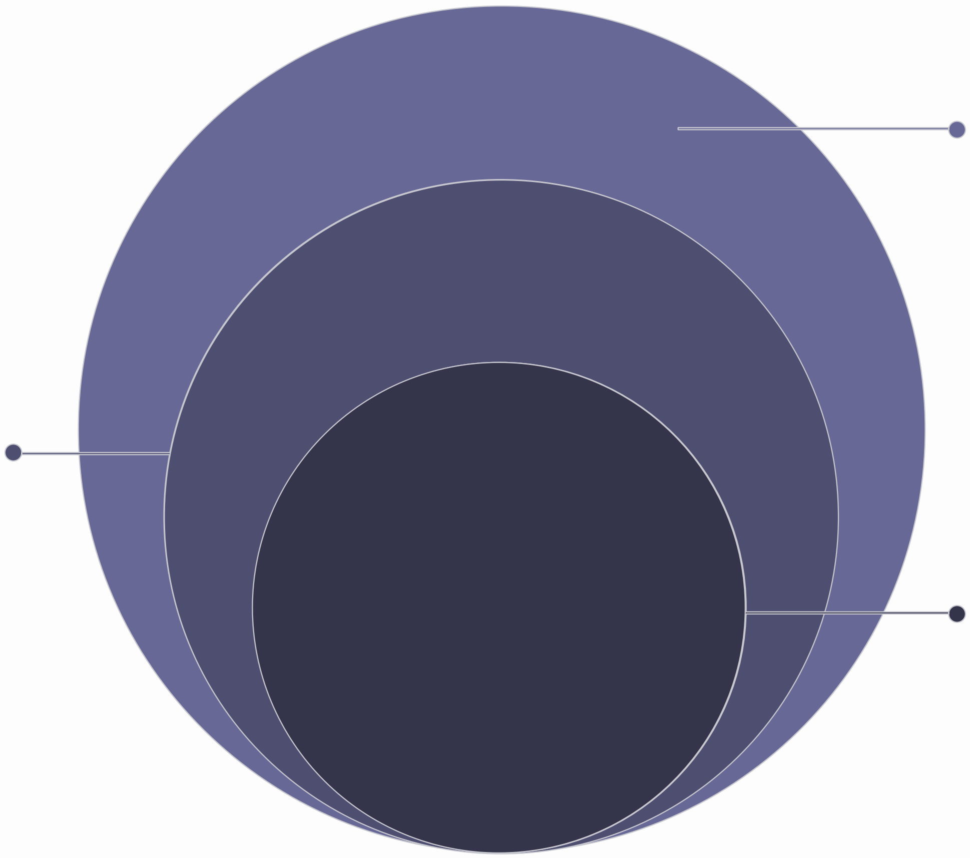
Severity Assignment
Risk-based prioritization

Analytics Dashboard
Actionable compliance intelligence

Core Innovation: We convert compliance documents into executable logic, enabling continuous automated enforcement instead of periodic manual audits.

System Architecture

The platform follows a modular, layered architecture designed for enterprise scalability and maintainability. Each layer maintains clear separation of concerns while enabling seamless data flow from user interaction to compliance enforcement.



Interface & Ops

Frontend, API, scheduler, and monitoring

Processing Engines

AI extraction and violation detection

Core Database

PostgreSQL stores rules and events

Architecture Principles

- **Modularity:** Independent components enable parallel development and isolated testing
- **Scalability:** Horizontal scaling at each layer supports enterprise transaction volumes
- **Security:** Defense-in-depth with authentication, encryption, and access controls
- **Resilience:** Fault isolation prevents cascading failures across layers

Technology Choices

Each layer uses purpose-built technologies optimized for its specific requirements—React for responsive UI, Node.js/FastAPI for scalable APIs, PostgreSQL for ACID compliance, and scheduled workers for reliable automation.

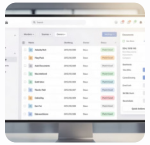
This design supports cloud deployment patterns including containerization, service mesh integration, and multi-region distribution.

Frontend Architecture

Built with React, the frontend delivers real-time compliance visibility through a component-based architecture optimized for performance and user experience. The modular design enables rapid feature development while maintaining code quality.

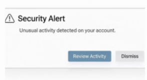
Dashboard Module

Real-time violation metrics, trend analysis, and severity breakdowns with interactive filtering and drill-down capabilities.



Policies Module

Policy upload, version management, rule preview, and status tracking with change history and rollback support.



Security Alert

Unusual activity detected on your account.

Violations Module

Searchable violation list with severity filters, status management, and detailed inspection views for investigation.

Settings & Monitoring

Database connection configuration, scan scheduling, notification preferences, and system health monitoring.

Technical Stack Benefits

- Component reusability reduces development time
- Virtual DOM enables fast state updates for real-time data
- Centralized state management ensures data consistency
- Responsive design adapts to desktop and mobile contexts

API Integration

The frontend communicates exclusively through RESTful APIs, maintaining clean separation from backend logic. Optimistic updates provide immediate user feedback while background synchronization ensures data accuracy.

Backend API Architecture

The backend API layer serves as the critical orchestration hub, providing secure, scalable endpoints that coordinate between the frontend, AI engine, database, and monitoring systems. Built with Node.js/FastAPI, it implements enterprise-grade patterns for authentication, rate limiting, and error handling.

1

File Upload API

Securely ingests policy PDFs with virus scanning, size validation, and encrypted storage. Generates unique identifiers and maintains upload audit logs.

2

Rule Extraction API

Orchestrates AI processing pipeline by sending parsed document content to the extraction engine, managing job queues, and handling asynchronous processing callbacks.

3

Database Connection API

Establishes secure PostgreSQL connections with credential encryption, connection pooling, and health checks. Supports multiple database configurations for multi-tenant scenarios.

4

Scan Trigger API

Initiates on-demand compliance scans with configurable scope (full database, date range, specific policies). Returns job ID for status tracking.

1

Scheduled Scan API

Configures automated periodic scans using cron expressions. Manages scheduling queue, retry logic, and failure notifications for reliable automation.

2

Violations Fetch API

Returns paginated violation results with advanced filtering (severity, status, date range, policy ID). Supports sorting and field selection for optimized payload size.

3

Dashboard Analytics API

Aggregates compliance metrics across configurable time windows. Pre-computes trends and severity distributions to minimize frontend computation and ensure fast load times.

Why REST APIs? Decoupled architecture enables independent scaling of services, simplifies cloud deployment, supports multiple client types, and accelerates development through clear interface contracts.

AI Rule Extraction Engine

The AI engine transforms unstructured compliance language into executable rule logic through a multi-stage NLP pipeline. This is where regulatory prose becomes programmatic enforcement—the core innovation that eliminates manual rule coding.

Processing Pipeline

01

PDF Parsing

Extract text while preserving document structure, tables, and hierarchical organization using OCR-enhanced parsing.

02

NLP Analysis

Identify rule patterns, extract conditional logic, detect thresholds, and recognize severity indicators through domain-specific language models.

03

Prompt Engineering

Structured prompts guide AI to generate consistent rule schemas with proper typing, validation logic, and edge case handling.

04

JSON Generation

Output validated structured rules ready for database insertion and immediate enforcement execution.

Why AI is Essential

Legal and regulatory language is inherently unstructured, filled with contextual nuance that resists traditional parsing. Manual rule writing by engineers takes days per document and requires legal expertise.

AI accelerates this process from days to minutes while maintaining accuracy through validation layers and human-in-the-loop review workflows.

```
{
  "rule_id": "AML_001",
  "condition": "transaction_amount > 10000 AND country IN ['high_risk']",
  "threshold": 10000,
  "severity": "HIGH",
  "action": "FLAG_FOR_REVIEW",
  "description": "Large transactions from high-risk jurisdictions"
}
```

The structured output enables immediate programmatic evaluation against transaction data, creating a continuous feedback loop between policy updates and enforcement.

Database & Detection Logic

PostgreSQL serves as the system's source of truth, storing policies, rules, transactions, and violation records with full ACID compliance. The detection engine evaluates transactions against rule conditions with millisecond-level performance at enterprise scale.

Why PostgreSQL?

- **ACID Guarantees:** Critical for compliance audit trails and regulatory reporting
- **Complex Queries:** Native support for JSON, full-text search, and window functions
- **Reliability:** Proven enterprise stability with point-in-time recovery
- **Extensibility:** Custom functions, triggers, and stored procedures for business logic

Schema Design

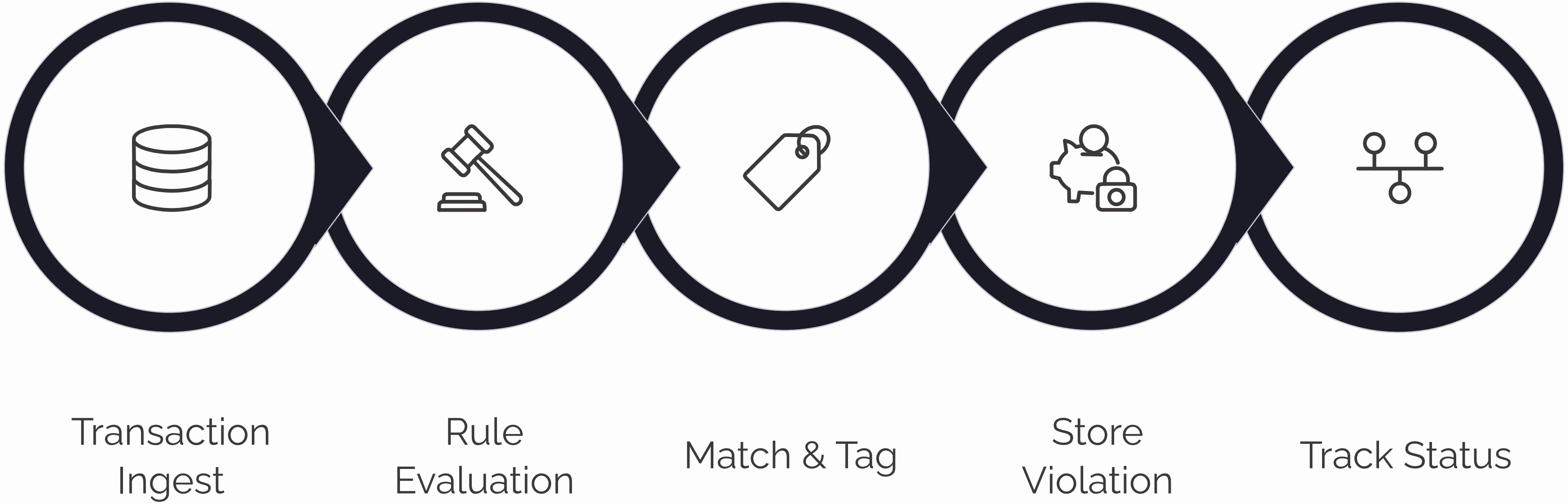
policies: Document metadata, version history, activation status

rules: Extracted conditions, thresholds, severity mappings

transactions: Customer data integration point (read-only views)

violations: Detected issues with severity, status, timestamps

scan_logs: Audit trail of all scanning activity and outcomes



Transaction Ingest

Rule Evaluation

Match & Tag

Store Violation

Track Status

The detection engine uses optimized SQL queries with indexed lookups to evaluate millions of transactions per hour. Each violation is tagged with severity level (Critical, High, Medium, Low) based on rule configuration and stored with complete context for investigation including triggering transaction details, applicable rule, and detection timestamp.

Monitoring, Automation & Analytics

The platform supports both on-demand and automated compliance scanning through flexible scheduling capabilities. Continuous monitoring ensures violations are detected within minutes of occurrence rather than weeks later during manual audits.

Manual Scan

Triggered on-demand for immediate compliance checks after policy updates or during investigations. Supports scoped scanning of specific time ranges or transaction subsets.

Scheduled Automation

Cron-based scheduling enables hourly, daily, or custom interval scans. Background job queues handle processing with retry logic and failure alerting.

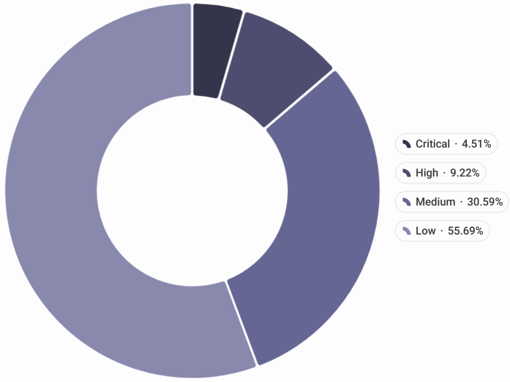
Continuous Monitoring

Real-time transaction stream evaluation for high-risk scenarios. Event-driven architecture enables sub-second violation detection and notification.

Audit Logging

Complete audit trail of all scanning activity, rule changes, and system events for regulatory compliance and forensic analysis.

Dashboard Intelligence



Real-Time Metrics: Total violations, severity distribution, resolution rates, and compliance score trending

Temporal Analysis: 7/14/30/90-day trend analysis identifies emerging patterns and systemic issues

Pending vs Resolved: Track remediation progress and measure team response effectiveness

Visual analytics transform raw violation data into actionable intelligence for risk officers and compliance teams, enabling data-driven prioritization and resource allocation.

Impact, Scalability & Future Vision

95%

Audit Time Reduction

Automated scanning eliminates manual spreadsheet reviews

60%

Cost Savings

Reduced compliance staff requirements and penalty avoidance

100x

Detection Speed

Real-time identification vs weekly/monthly manual audits

Scalability Architecture

- **Multi-Tenant SaaS:** Isolated data, shared infrastructure for enterprise deployment
- **Cloud Native:** Container orchestration, auto-scaling, multi-region support
- **Database Flexibility:** MySQL, MSSQL, Oracle integration through abstraction layer
- **High Availability:** Redundant services, failover mechanisms, zero-downtime updates

Future Roadmap

Multi-Model AI: Ensemble approaches for improved extraction accuracy across regulatory domains

Streaming Detection: Kafka integration for real-time transaction evaluation at massive scale

Compliance-as-a-Service: API-first platform enabling embedded compliance for fintech ecosystems

Compliance should not be reactive. It should be autonomous and continuous.

Data Policy Agent transforms compliance from a periodic manual burden into an automated, intelligent system that operates 24/7. By converting regulatory documents into executable logic, we enable organizations to detect violations in real-time, reduce risk exposure, and build trust with regulators and stakeholders. The future of compliance is proactive, data-driven, and powered by AI.