# AI-powered system that automates data validation, compliance rule extraction, and remediation suggestions

Manual data validation and compliance checks are time-consuming, error-prone, and struggle to keep up with evolving regulations. Organizations face challenges in extracting complex compliance rules and applying them consistently. Lack of automated remediation suggestions leads to inefficiencies and compliance risks. An AI-driven solution is needed to streamline validation, extract rules, and provide actionable remediation insights.

Our Solution - AI-Powered Compliance System

What does our solution do?

✅ Extracts validation rules from regulatory documents

✅ Generates automated validation logic for data files

✅ Detects compliance issues and anomalies

✅ Provides AI-driven remediation recommendations

✅ Creates audit-ready reports

Key Technologies Used:

• NLP (Large Language Models)

• Machine Learning (Rule Generation & Anomaly Detection)

• Python (Validation Code Generator)

• Great Expectations (Data Validation Framework)

• Vector Embeddings (Regulatory Document Processing)

System Architecture (Layered Approach)

Input Sources Layer

What it does: Collects emails, documents (PDFs, CSVs, TXT), and regulatory guidelines

Converts unstructured text into a structured format for processing

AI Orchestration Layer

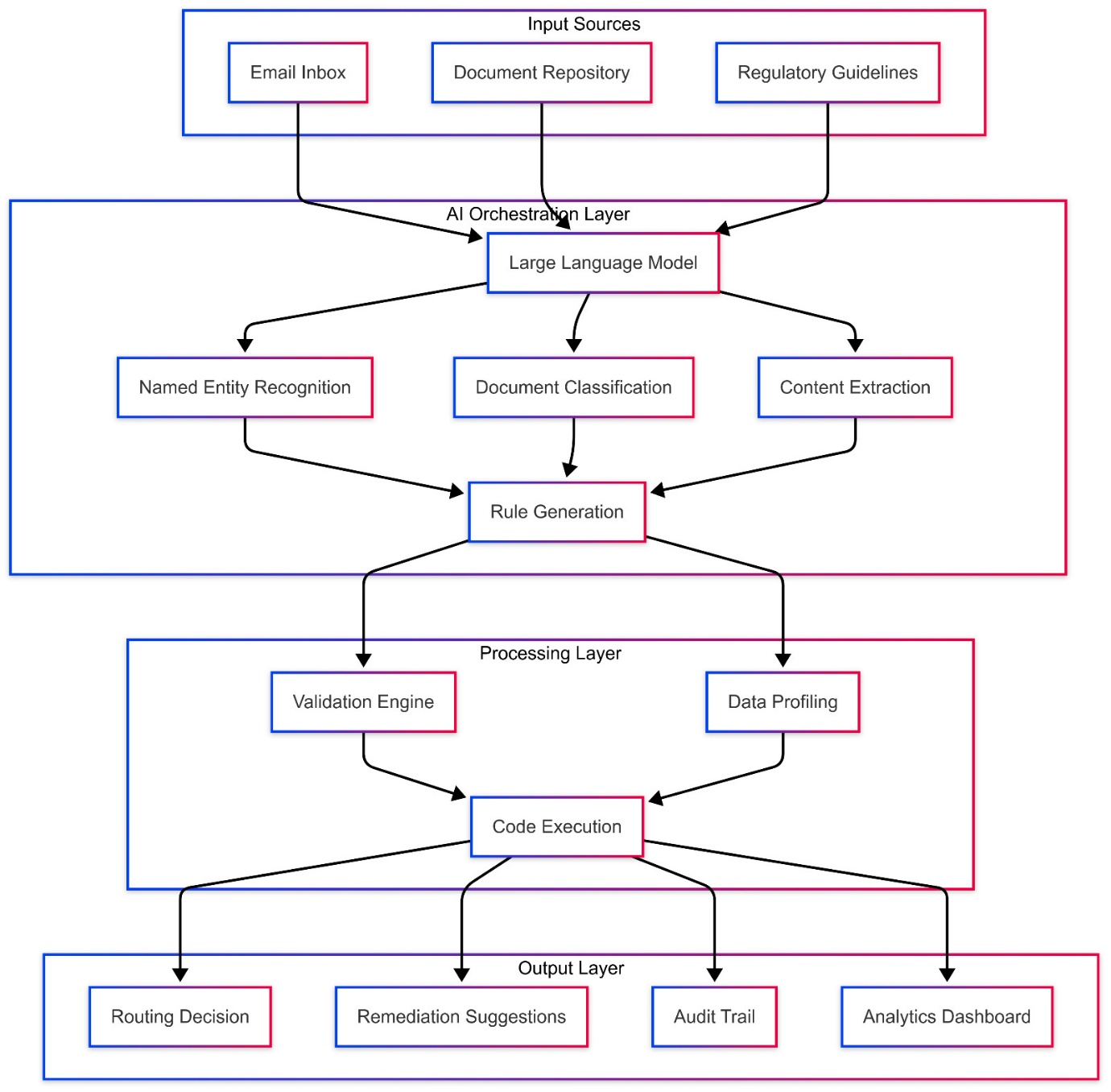
Uses LLMs & NLP to extract compliance rules

Identifies key validation criteria: Required fields, ranges, cross-checks

Processing Layer

Data Profiling & Validation

Identifies missing values, incorrect formats, rule violations



## Key System Components

Regulatory Instruction Processor

🔹 Reads compliance documents & extracts rules

🔹 Uses vector embeddings for efficient document search

🔹 Identifies required fields, data types, range checks

Rule Generator

🔹 Discovers new validation rules using unsupervised learning

🔹 Identifies categorical values, statistical patterns, and correlations

🔹 Converts extracted rules into machine-executable validation templates

Validation Code Generator

🔹 Generates Python-based validation scripts

🔹 Creates Great Expectations validation suites for standard checks

🔹 Implements rule-based range checks, null checks, cross-field validations

Remediation Recommender

🔹 Analyzes failed validations, suggests fixes & explanations

🔹 Determines which issues can be auto-fixed vs. manual review

🔹 Generates audit-ready compliance explanations

Document Processor

🔹 Extracts key data from emails, PDFs, and CSVs

🔹 Classifies documents based on type and content

🔹 Determines appropriate routing for compliance processing

Main Orchestrator

🔹 Manages the entire compliance workflow

🔹 Processes regulatory documents, validates data, and applies remediations

🔹 Generates audit-ready compliance reports

Dashboard UI

🔹 Displays validation results, compliance status, and remediation actions

🔹 Provides real-time monitoring of compliance issues

## End-to-End Workflow Example

* Step 1: Input Processing
  + System receives regulatory documents, emails, and structured data (CSV, Excel, etc.)
* Step 2: AI-Driven Processing
  + Extracts compliance rules
  + Generates validation logic
  + Runs validation checks
  + Identifies missing fields and anomalies
* Step 3: Output & Compliance Actions
  + Validated Data
  + Audit Reports
  + Automated Fixes & Remediation Plans

## Dashboard UI & Reporting

* Validation Status Reports – Shows compliance check results
* Rule-based Errors & Fixes – Identifies violations & recommends corrections
* Audit Logs – Tracks all automated and manual changes

## Why This Solution is a Game-Changer?

* Reduces Manual Effort – Automates 80% of compliance validation
* Enhances Accuracy – AI-based validation reduces errors
* Improves Efficiency – Fast rule processing & automatic fixes
* Ensures Regulatory Compliance – Audit-ready reports & rule tracking
* Potential Impact:
* Saves 100s of manual hours for compliance teams
* Reduces regulatory penalties by ensuring data accuracy
* Can be integrated into real-world banking & financial systems.

### Next Steps & Roadmap

* Phase 1: Build MVP with rule extraction & validation
* Phase 2: AI-driven anomaly detection & auto-fixes
* Phase 3: Dashboard with real-time compliance monitoring