ESP IJAST
ESP International Journal of Advancements in Science & Technology
ISSN: 2583-9233 / Volume 3 Issue 4 December 2025 / Page No: 13–16
Paper Id: IJAST-V3I4P103 / Doi: 10.56472/25839233/IJAST-V3I4P103

Original Article

AI for Regulated Domains: Bridging Compliance, Explainability, and Trust

Beng Ee Lim[1], Mert Zamir[2]
Independent Researcher.

Received Date: 15 November 2025 Revised Date: 29 November 2025 Accepted Date: 31 December 2025

Abstract: *Artificial intelligence is increasingly used to support complex decision workflows, yet its adoption in regulated industries remains limited. Medical device regulation, pharmaceutical oversight, aviation safety, and financial governance all rely on strict documentation, traceable reasoning, and verifiable evidence. Current large language models struggle to meet these requirements because they do not natively provide citations, cannot guarantee consistency, and lack the mechanisms needed for repeatable reasoning. This paper outlines the core challenges of applying AI in tightly regulated domains, identifies the gaps in current AI-assisted workflows, and proposes a practical agentic framework that emphasizes explainability, traceability, and trust. The goal is to show how AI can enhance compliance work without undermining regulatory integrity.*

Keywords: *Artificial Intelligence, Regulated Domains, Explainable AI (XAI),Trustworthy AI, Agentic AI, Regulatory Compliance, Auditability and Traceability.*

I. INTRODUCTION

Many industries operate in environments where mistakes carry real consequences. Medical device manufacturers must satisfy FDA guidelines before entering the market. Aviation regulators impose strict certification standards for airworthiness. Banks are required to justify risk models to supervisory authorities. These processes depend on precise documentation, careful interpretation of rules, and the ability to defend every decision made.

AI has shown promise in accelerating document search, summarizing technical information, and identifying patterns in large datasets. Yet adoption remains cautious. Regulators, compliance officers, and risk managers share similar concerns: Can the system explain its output? Can the source be verified? Can the reasoning be reproduced?

This paper examines why conventional AI systems struggle in these environments and presents a structured approach for building agentic AI workflows that respect regulatory expectations. The discussion is grounded in the medical device sector, though the principles extend to any domain where decisions must be auditable.

II. BACKGROUND: WHY REGULATED DOMAINS ARE DIFFICULT FOR AI

Regulated work differs from typical enterprise tasks in several ways:

- First, rules change slowly but accumulate over decades. Guidance documents, standards, and public notices often span thousands of pages. Many contain edge cases, exceptions, and historical clauses that remain relevant.
- Second, artifacts are diverse in structure. Regulations appear as PDFs, HTML pages, scanned documents, and structured datasets. These formats vary across agencies and time periods.
- Third, decisions require evidence. It is not enough to produce a correct answer. A compliance engineer must point to the exact paragraph, table, or precedent supporting each claim.
- Fourth, ambiguity is treated differently. In software development, ambiguity can be resolved through assumptions. In regulated work, assumptions—if unstated—can invalidate a submission.

This combination of complexity, heterogeneity, and accountability creates a setting where AI must behave differently from typical consumer or enterprise applications.

III. LIMITATIONS OF CURRENT AI APPROACHES IN COMPLIANCE TASKS

Despite rapid progress, mainstream AI systems share several shortcomings that limit their use in regulatory settings.

A. Absence of Verifiable Citations

Most models generate fluent text but cannot reliably point to the source that supports a claim. Even when retrieval is used, citations may be incomplete or inconsistent.

B. Inconsistent Reasoning

Given the same prompt twice, models may produce different answers. In high-stakes contexts, reproducibility is essential.

C. Difficulty Understanding Structured Rules

Regulatory documents often include nested clauses, conditional logic, and cross-referenced sections. Current models struggle to track these relationships without explicit prompting.

D. Sensitivity to Incomplete Context

If the model receives an incomplete or poorly retrieved set of documents, it will still produce an answer. This creates risk when the system appears confident but lacks the necessary evidence.

These limitations show that conventional LLMs are not sufficient on their own. Additional structure is needed to support traceability and trust.

IV. REQUIREMENTS FOR TRUSTWORTHY AI IN REGULATED DOMAINS

Through industry practice and preliminary deployments, several requirements emerge as consistent expectations among regulators and compliance professionals.

A. Grounded Retrieval

Systems must retrieve only documents that are relevant and valid. This includes ensuring that outdated or superseded regulations are excluded.

B. Clear Source Attribution

Every output should reference a precise section of the document on which it relies. These references must be visible and accessible.

C. Transparent Reasoning Steps

Compliance professionals expect to understand how information was interpreted. An AI system should outline its reasoning in a structured manner, allowing review of each step.

D. Support for Human Verification

AI should reduce manual effort, not eliminate oversight. Users must be able to validate, correct, and refine outputs easily.

E. Auditability

For future inspections or internal reviews, the system must log which documents were used, how they were retrieved, and how the final answer was produced.

These principles guide the development of agentic AI workflows suited for regulatory applications.

V. AGENTIC AI FRAMEWORK FOR REGULATED WORKFLOWS

The paper proposes an agentic architecture built on four layers: retrieval, interpretation, planning, and verification. Each layer addresses a specific failure mode of traditional AI systems.

A. Retrieval Layer

This layer collects the relevant regulatory documents. It prioritizes structured ingestion, deduplication, and version control. By ensuring that the system works with up-to-date information, retrieval reduces downstream errors.

B. Interpretation Layer

Here, the AI parses the retrieved documents and identifies the sections that may be relevant to the user's question. The system presents these excerpts along with the context needed for human review.

C. Planning Layer

Agentic models break down complex regulatory questions into smaller steps. For example, determining the testing requirements for a device may involve identifying the product code, matching it to applicable standards, and reviewing historical clearances.

D. Verification Layer

This layer ensures that each step is accompanied by supporting citations. It checks for missing references, inconsistent logic, or unsupported conclusions, then provides a summary that highlights any issues needing human review.

This layered structure helps align AI behavior with regulatory expectations.

VI. SYSTEM ARCHITECTURE OVERVIEW

A typical implementation includes:

A. Document Ingestion Pipeline

PDFs, JSON datasets, and HTML pages are processed through text extraction or OCR. Metadata is standardized. Outdated documents are tagged or removed.

B. Embedding and Semantic Indexing

Documents are chunked in a way that preserves logical structure. Semantic embeddings are created to support precise retrieval.

C. Agentic Controller

The controller receives user queries, plans the required steps, and calls retrieval or analysis tools as needed.

D. Traceability Engine

Each output is linked to its originating source. Users can inspect the sections that informed each line of reasoning.

E. Human-in-the-Loop Interface

Compliance officers can review the AI's reasoning, accept or reject steps, and add clarifications. This interaction creates a shared trail of work for future audits.

VII. EXAMPLE SCENARIO: FDA MEDICAL DEVICE REGULATION

To illustrate the framework, consider a company preparing a submission for a Class II medical device. One of the early tasks is to determine the applicable standards and testing requirements.

When a user asks the system to identify the relevant testing pathways, the AI performs several steps:

- Retrieves the product classification and associated guidance documents.
- Extracts the testing requirements written in each document.
- Identifies any conflicting or outdated standards.
- Provides excerpts that explain the reasoning.
- Offers a structured summary that the regulatory team can validate.

This workflow helps users interpret complex documentation without losing sight of the evidence. It also reduces the risk of overlooking an important requirement.

VIII. DISCUSSION: BUILDING TRUST IN AI-SUPPORTED COMPLIANCE

Trust emerges when an AI system behaves predictably and transparently. Regulators are not opposed to AI involvement, but they expect systems to operate with clear boundaries.

Several observations arise from early deployments:

- Users trust systems that reveal uncertainty. Attempts to hide uncertainty reduce confidence.
- Consistency matters more than speed. Fast answers without citations are not useful in regulated settings.
- Humans prefer systems that explain their steps. Even brief outlines of reasoning build trust.
- Version awareness is essential. New standards and guidance documents appear regularly, and the system must track these changes.

These findings suggest that trust is earned through careful system design rather than advanced model performance alone.

IX. LIMITATIONS AND FUTURE DIRECTIONS

Although promising, agentic AI for regulated domains faces several limitations.

- Complex legal language remains difficult to parse. Some documents include ambiguous phrasing or multi-layered references.
- Interpretation still requires expert input. AI cannot replace professional judgment.
- Validation frameworks are still evolving. Standards for evaluating AI-generated reasoning in compliance contexts are under development.
- Cross-agency differences complicate generalization. Each regulator follows different conventions.

Future work may involve integrating formal verification techniques, developing domain-specific training corpora, and creating shared benchmarks for evaluating agentic AI in regulated settings.

X. CONCLUSION

Regulated industries demand precise, transparent, and defensible decisions. Traditional AI models excel at generating fluent text, yet they fall short in tasks that require verifiable evidence and consistent reasoning. Agentic AI offers a way to bridge this gap by structuring the workflow into retrieval, interpretation, planning, and verification. Through this layered approach, AI can support compliance professionals while maintaining the standards of accountability required in safety-critical fields.

As AI systems mature, the focus should shift from raw model performance to frameworks that emphasize trust and reliability. This shift will enable responsible innovation in domains where accuracy and traceability are essential.