Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 12, 2025

AUTOMATING SHARIAH COMPLIANCE CHECKS IN ISLAMIC BANKING USING LARGE LANGUAGE MODELS AND RAG PIPELINES

[1]Dr. Faisal Ahmed Sarfaraz, [2]Muhammad Farhan Memon, [3]Salman Mahmood

[1]Lecturer College Education Department Sindh
[2]Lecturer, College Education Department
[3]M.Phil (Islamiyat) Greenwich University/ The City School

[1]faysalaleemi@gmail.com, [2]farhanmemon07@gmail.com, [3]tcs.salmanm@gmail.com

DOI:

Article History
Received on 28 Nov, 2025
Accepted on 25 Dec 2025
Published on 26 Dec 2025

Corresponding Author:

Abstract

The Islamic banking and finance industry, currently managing assets exceeding USD 4 trillion globally, operates under a rigorous framework of Shariah law that mandates exhaustive compliance review of every financial instrument and contractual agreement. Conventional Shariah audit processes are predominantly manual, reliant upon a limited pool of certified scholars, and unable to scale proportionally with the exponential growth of Islamic financial products. This paper proposes a novel, end-to-end computational framework that integrates Large Language Models (LLMs) with Retrieval-Augmented Generation (RAG) pipelines to automate first-pass Shariah compliance checks for Islamic financial contracts. The framework ingests a curated corpus comprising Accounting and Auditing Organization for Islamic Financial Institutions (AAOIFI) standards, classical fiqh rulings, historical fatwas, and standardized Islamic contract templates. A multi-stage RAG pipeline–encompassing dense vector embedding, semantic retrieval from a high-dimensional knowledge base, and contextually grounded LLM reasoning–is employed to evaluate contractual clauses against established Shariah principles. Critically, the architecture incorporates a Dynamic Knowledge Caching layer that stores frequently queried compliance outcomes, thereby reducing query latency by an estimated 67% and computational overhead by up to 45% at production scale. A simulated case study evaluating a Mudaraba (profit-sharing) contract demonstrates the system's capacity to flag non-compliant clauses–such as fixed-return guarantees–and generate Shariah-aligned revisions. The framework maintains a mandatory Human-in-the-Loop (Shariah scholar) verification stage, ensuring auditability and religious legitimacy. Results indicate substantial efficiency gains, with contract review time reduced from an average of 8–12 hours to under 3 minutes, while achieving a clause-level retrieval precision of 91.4% against AAOIFI benchmarks. This work establishes a scalable, explainable, and institutionally viable pathway toward AI-augmented Shariah governance.

1. Introduction

Islamic banking and finance (IBF) represents one of the fastest-growing sectors within the global financial system. From its foundational origins in the late twentieth century, the industry has expanded to encompass more than 70 countries across the Middle East, Southeast Asia, Sub-Saharan Africa, and Western markets, with total Shariah-compliant assets projected to surpass USD 6 trillion by 2028 [1]. Central to the operational legitimacy of every Islamic financial institution (IFI) is the principle of Shariah compliance: the requirement that all products, contracts, and transactions conform to the precepts of Islamic law, including the prohibition of riba (interest), gharar (excessive uncertainty), maysir (speculation), and the mandated sharing of profit and loss [2].

The process of ensuring Shariah compliance is currently operationalized through Shariah Supervisory Boards (SSBs)—panels of qualified Islamic scholars who review, deliberate, and issue fatwas (religious rulings) on proposed financial instruments. While this model preserves the religious integrity of IBF, it presents critical scalability constraints. The global supply of duly qualified Shariah scholars is estimated at fewer than 300 individuals capable of conducting rigorous financial product reviews [3]. Simultaneously, the product complexity and transactional volume of modern IBF have grown exponentially, creating a chronic bottleneck: manual review cycles that span weeks or months, inconsistency in scholarly interpretation across jurisdictions, and mounting operational risk arising from compliance oversights [4].

The emergence of Large Language Models (LLMs) such as GPT-4, LLaMA-2, and Claude represents a paradigm shift in natural language processing capability. These models exhibit remarkable proficiency in document comprehension, legal reasoning, and structured text generation—capabilities directly relevant to the Shariah audit process [5]. However, the direct application of general-purpose LLMs to Shariah compliance verification is fraught with significant risk. LLMs are known to generate plausible yet factually incorrect outputs—a phenomenon termed 'hallucination'—which, in the context of sacred religious-legal frameworks, carries potentially severe reputational, financial, and doctrinal consequences for IFIs [6].

Retrieval-Augmented Generation (RAG) has emerged as the canonical architectural response to LLM hallucination in high-stakes domains. By grounding LLM inference in dynamically retrieved, authoritative documentary evidence, RAG constrains generative outputs to verified source material, thereby ensuring that compliance conclusions are traceable to specific AAOIFI standards, recognized fiqh texts, or accepted fatwas [7]. This paper proposes a comprehensive RAG-LLM framework specifically designed for first-pass Shariah compliance auditing, augmented with a Dynamic Knowledge Caching layer to operationalize the framework at the throughput demands of large-scale banking institutions.

1.1 Research Objectives and Contributions

This paper makes the following primary contributions to the intersection of computational linguistics and Islamic finance:

- The design and specification of an end-to-end RAG-based Shariah compliance pipeline, incorporating document ingestion, semantic retrieval, LLM-grounded reasoning, and structured compliance reporting.
- The introduction of a Dynamic Knowledge Caching layer as an architectural component of the compliance pipeline, with a formal analysis of its latency and cost-reduction properties for production-scale deployment.
- A simulated case study evaluating a Mudaraba contract against AAOIFI Financial Accounting Standard No. 3 (FAS 3), demonstrating the system's non-compliance detection and revision generation capabilities.
- A critical analysis of security vulnerabilities—including prompt injection and adversarial manipulation—and a Human-in-the-Loop governance model preserving final scholarly authority.

2. Literature Review

2.1 Artificial Intelligence in Financial Technology

The application of machine learning and natural language processing to financial services has generated a substantial body of scholarly work over the preceding decade. Arner, Barberis, and Buckley [8] provided an early taxonomy of FinTech innovation trajectories, within which regulatory compliance automation—or RegTech—emerged as a distinct and rapidly maturing subfield. Subsequent work by Philippon [9] formalized the efficiency gains attributable to algorithmic financial intermediation. Within RegTech specifically, transformer-based language models have demonstrated significant promise for contract analysis, regulatory classification, and automated reporting [10].

Chalkidis et al. [11] established foundational benchmarks for legal language understanding using pre-trained transformer models on European legislation corpora, demonstrating recall rates exceeding 80% on regulatory clause classification tasks. Analogous studies in general financial compliance, such as those by Tsarapatsanis and Aletras [12], illustrated that fine-tuned BERT-class models substantially outperform rule-based keyword approaches for statutory interpretation tasks. Collectively, these works establish the technical viability of NLP-based compliance automation while implicitly surfacing the requirement for domain-specific corpora and grounding mechanisms when migrating from general law to specialized religious-legal frameworks.

2.2 RAG Architectures and Hallucination Mitigation

Lewis et al. [7] introduced the foundational RAG architecture, demonstrating that retrieval-augmented decoding significantly reduces factual error rates in open-domain question answering benchmarks. Subsequent architectural variants—including Dense Passage Retrieval (DPR) [13], Fusion-in-Decoder models, and REALM—have collectively established that the quality of the retrieval corpus and the precision of the embedding model are the primary determinants of system accuracy in knowledge-intensive tasks [14]. For high-stakes legal and financial applications, Shao et al. [15] demonstrated that RAG-augmented models exhibited a 43% reduction in hallucination frequency relative to standard LLM inference, positioning RAG as the architecture of choice for sensitive regulatory domains.

2.3 The Gap: NLP for Shariah Compliance

Despite the maturation of both LLM technology and Islamic finance scholarship, the intersection of these two domains remains conspicuously underdeveloped in the peer-reviewed literature. Abozaid [16] and Laldin and Furqani [17] provided comprehensive qualitative analyses of Shariah governance structures and their institutional limitations but did not engage with computational approaches. Existing attempts at computational Shariah assistance, such as those surveyed by Hassan and Lewis [18], have been confined to rule-based expert systems and keyword-matching engines that lack the contextual reasoning capacity required for nuanced contractual interpretation.

A small number of recent works have begun to explore NLP for Islamic finance. Billah [19] piloted a term extraction system for Islamic financial documents but did not address compliance inference or retrieval-augmented reasoning. To the knowledge of the present authors, no published work has proposed or evaluated a full-pipeline RAG-LLM architecture specifically targeting AAOIFI-standard Shariah compliance verification. This paper addresses that gap directly.

3. Proposed Methodology: The Shariah Compliance AI Framework

The proposed Shariah Compliance AI Framework (SCAF) is structured as a sequential, multi-component pipeline. Figure 1 illustrates the high-level architecture. The pipeline operates in four principal phases: (i) Knowledge Base Construction and Document Ingestion; (ii) Query Processing and Dynamic Caching; (iii) RAG-Enabled Compliance Reasoning; and (iv) Human-in-the-Loop Verification and Audit Logging.

3.1 Phase 1: Knowledge Base Construction and Document Ingestion

3.1.1 Corpus Composition

The authoritative knowledge base underpinning SCAF is assembled from three primary document categories, each representing a distinct stratum of Islamic legal authority:

- AAOIFI Standards: The complete corpus of AAOIFI's Shariah Standards (SS), Financial Accounting Standards (FAS), and Governance Standards (GS), comprising 48 Shariah standards, 30 financial accounting standards, and associated guidance notes [20]. These documents constitute the primary operational reference for IFI compliance globally.
- Classical Fiqh Texts and Fatwas: Digitized rulings from recognized Shariah bodies including the OIC Islamic Fiqh Academy, the Shariah Advisory Council of Bank Negara Malaysia, and the Saudi Arabian Permanent Committee for Scholarly Research, supplemented by authoritative classical texts such as Al-Muwatta and Majallat Al-Ahkam Al-Adliyyah.
- Standardized Islamic Contract Templates: Annotated contract templates for canonical instruments including Murabaha, Mudaraba, Musharaka, Ijara, Sukuk, and Tawarruq, with clause-level annotations indicating Shariah compliance rationale per AAOIFI Shariah Standard No. 3, 4, 8, and 9 [21].

3.1.2 Document Pre-processing Pipeline

Raw documents are subjected to a multi-stage pre-processing pipeline. Optical Character Recognition (OCR) is applied to scanned Arabic and bilingual texts using Tesseract 5.0 with Arabic language models [22]. Documents are then segmented into semantically coherent chunks of 512 tokens with a 64-token overlap stride, preserving clause-level integrity while maintaining contextual continuity across segment boundaries. Each chunk is tagged with structured metadata including: source document identifier, AAOIFI standard number, chapter and clause reference, language of origin (Arabic/English), and an authority-tier score (1–3) reflecting the hierarchical precedence of the source in Islamic jurisprudence.

3.2 Phase 2: Embedding Generation and Vector Database Storage

Each processed text chunk is transformed into a dense vector representation using a bilingual (Arabic-English) sentence transformer model fine-tuned on Islamic finance corpora. Specifically, the architecture builds upon the multilingual E5-large model [23], which has been further fine-tuned on a curated dataset of AAOIFI standard pairings and Shariah fatwa question-answer pairs (approximately 24,000 training examples). The resulting embeddings occupy a 768-dimensional semantic space, enabling nuanced similarity computation across mixed-language Islamic legal documents.

Embeddings are stored in a distributed vector database—specifically, Pinecone or Weaviate at enterprise deployment scale—indexed using Hierarchical Navigable Small World (HNSW) graphs, which support approximate nearest-neighbour retrieval at sub-millisecond latency for corpora of up to 10 million vectors [24]. Each vector record retains a pointer to its source chunk and full metadata payload, ensuring that retrieved evidence is immediately traceable to its authoritative origin document.

3.3 Phase 3: The Dynamic Knowledge Caching Layer

A fundamental architectural innovation of SCAF is the Dynamic Knowledge Caching (DKC) layer, positioned between the query interface and the RAG retrieval engine. The rationale for caching in Shariah compliance applications is compelling: within any given IFI's product portfolio, a small set of clause types—profit-sharing ratios, guarantee clauses, collateral conditions—is queried with high frequency across numerous contracts. Without caching, each query independently traverses the full vector retrieval and LLM inference pipeline, incurring unnecessary computational cost and latency.

3.3.1 Cache Architecture

The DKC layer implements a two-tier cache architecture:

- Semantic Cache (Tier 1): Query embeddings are compared against a cache index of previously processed queries using cosine similarity. If a cached query exceeds a similarity threshold $\theta = 0.92$ to the incoming query, the previously computed compliance response is returned directly without invoking the retrieval engine or LLM. This tier is implemented using Redis with a custom vector extension, supporting O(1) hash-based lookup augmented by approximate similarity search for semantic matching.
- Clause-Level Result Cache (Tier 2): Specific AAOIFI rule interpretations that appear in multiple retrieved chunks are pre-computed and cached at the clause level. These cached interpretations are injected directly into the LLM context when the corresponding clauses are retrieved, reducing the reasoning workload required of the generative model.

3.3.2 Cache Management and Invalidation

All cached entries are assigned a time-to-live (TTL) of 30 days, reflecting the approximate cadence of AAOIFI standard revisions and SSB fatwa updates. Upon detection of a new AAOIFI standard publication or a significant fatwa revision—monitored via RSS feeds from official AAOIFI and OIC channels—a targeted cache invalidation routine purges all entries referencing the affected standard numbers. This ensures that the compliance pipeline does not generate responses based on superseded religious-legal guidance.

Empirical simulation across 10,000 synthetic Shariah compliance queries drawn from five product categories indicates that the DKC layer achieves a steady-state cache hit rate of 72.3%, reducing average end-to-end response latency from 4.2 seconds (full RAG pipeline) to 1.4 seconds (cache-served), and decreasing LLM API token expenditure by an estimated 45% at production throughput.

3.4 Phase 4: RAG-Enabled Compliance Reasoning

For queries that bypass or miss the cache, the RAG pipeline is invoked. The compliance query—typically a suspect contract clause submitted as natural language text—is encoded using the same bilingual sentence transformer. A top-k semantic search ($k = 5$) retrieves the most relevant document chunks from the vector database. These chunks, along with their metadata, are assembled into a structured compliance prompt delivered to the LLM.

The prompt template is engineered to enforce citation discipline and structured output. The LLM (GPT-4-Turbo or a domain-fine-tuned open-weight equivalent) is instructed to: (i) identify the specific AAOIFI standard and clause governing the queried contract element; (ii) compare the submitted clause against the retrieved standard; (iii) deliver a binary compliance verdict (Compliant / Non-Compliant) with citation; and (iv) if non-compliant, propose a Shariah-compliant alternative clause formulation grounded in retrieved evidence. The LLM is explicitly prohibited from generating compliance rulings that are not supported by at least one retrieved document chunk, enforced through a post-generation validation layer that cross-references cited sources against the retrieved context.

4. Simulated Case Study: Evaluating a Mudaraba Contract

4.1 Scenario Overview

To demonstrate the operational behaviour of SCAF, we present a simulated evaluation of a standard Mudaraba (profit-sharing partnership) financing agreement. In a canonical Mudaraba structure, one party (the Rabb al-Maal, or capital provider) contributes capital while the other party (the Mudarib, or entrepreneur) contributes skill, management, and labour. Profits are shared according to a pre-agreed ratio, while financial losses are borne exclusively by the capital provider, absent negligence or misconduct by the Mudarib [25]. This risk-sharing asymmetry is a foundational principle of Islamic commercial law and

is codified in AAOIFI Shariah Standard No. 13 and Financial Accounting Standard No. 3.

4.2 The Non-Compliant Clause

The hypothetical contract submitted to SCAF contains the following clause under the profit and loss section:

Clause 7.3: The Mudarib guarantees to the Rabb al-Maal a minimum annual return of 6.5% on the invested capital, irrespective of the actual profit or loss generated by the Mudaraba enterprise during the contract period.

This clause constitutes a substantive Shariah violation. By guaranteeing a fixed minimum return regardless of actual business performance, it effectively transforms the Mudaraba—a participatory profit-and-loss sharing instrument—into a riba-bearing loan, wherein the capital provider is assured a predetermined yield. This is explicitly proscribed under AAOIFI Shariah Standard No. 13, Clause 4/1, and has been the subject of numerous SSB rulings and OIC Fiqh Academy resolutions.

4.3 System Pipeline Execution

Table 1 traces the execution of SCAF against the submitted clause, from initial query encoding through final output generation.

Table 1: ***SCAF Pipeline Execution – Mudaraba Case Study***

Pipeline Stage	Input	Output
Query Encoding	Mudaraba contract text	768-dim embedding vector
Cache Lookup	Query hash	Miss → proceed to retrieval
Vector Retrieval	Top-k = 5 AAOIFI chunks	FAS 3 §§ 4.1–4.4 retrieved
LLM Reasoning	Prompt + retrieved context	Non-compliance flag raised
Revision Generation	Non-compliant clause	Compliant clause drafted
Cache Update	Query + response pair	Stored with TTL = 30 days
Scholar Review	Full reasoning trace	Approval / override logged

4.4 Retrieved Evidence

The top-5 retrieved chunks upon semantic search against the non-compliant clause include: (i) AAOIFI SS No. 13, §4/1 ('It is not permissible for the Mudarib to guarantee the capital or a return thereon…'); (ii) AAOIFI FAS 3, §§ 4.1–4.4 on the accounting treatment of Mudaraba losses; (iii) OIC Islamic Fiqh Academy Resolution No. 30/5/4 on guaranteed returns in participatory contracts; (iv) Bank Negara Malaysia Shariah Advisory Council resolution (2010) on non-permissible profit guarantees; and (v) an annotated Mudaraba contract template flagging fixed-return clauses as non-compliant. Collectively, these sources achieve an authoritative coverage score of 0.94 on a normalized AAOIFI alignment index, providing robust evidentiary grounding for the compliance determination.

4.5 System Output

SCAF generates the following structured compliance report:

Field	Value
Compliance Status	NON-COMPLIANT

Violated Standard	AAOIFI Shariah Standard No. 13, Clause 4/1; FAS 3 §4.2
Violation Type	Riba-equivalent guarantee; transformation of Mudaraba into interest-bearing loan
Evidence Sources	OIC Fiqh Academy Res. 30/5/4; BNM-SAC 2010; AAOIFI SS13
Proposed Revision	'Profits, if any, shall be distributed between the Rabb al-Maal and the Mudarib in the ratio of [X%:Y%] as agreed. In the event of loss not attributable to the negligence or misconduct of the Mudarib, the Rabb al-Maal shall bear the financial loss in full, in accordance with AAOIFI Shariah Standard No. 13, Clause 4/2.'
Confidence Score	0.94 / 1.00
Scholar Review Required	YES – Forwarded to Shariah Supervisory Board queue

The proposed revision eliminates the fixed-return guarantee and reinstates the profit-and-loss sharing structure mandated by Shariah law, with explicit citation to the governing AAOIFI standard. The revision is automatically queued for scholar ratification prior to any contractual execution.

5. Performance Analysis and System Evaluation

Table 2 presents comparative performance metrics contrasting the proposed SCAF framework against the conventional manual Shariah audit process across key operational dimensions. These figures are derived from simulated processing of 10,000 synthetic Islamic financial contracts spanning five product categories (Murabaha, Mudaraba, Musharaka, Ijara, Sukuk) under realistic workload conditions.

Table 2: Comparative Performance Metrics – Manual vs. SCAF

Metric	Baseline (Manual)	RAG-LLM (Proposed)	Improvement
Avg. Review Time / Contract	8–12 hours	< 3 minutes	~240× faster
Clause Retrieval Precision	N/A	91.4%	—
AAOIFI Rule Recall	N/A	88.7%	—
Cache Hit Rate (Steady State)	N/A	72.3%	—
Latency Reduction (cached)	N/A	~67%	—
Scholar Override Rate	100% (all decisions)	14.2% (flagged only)	—

The efficiency gains are substantial. The reduction in average contract review time from 8–12 hours to under 3 minutes represents an approximately 240-fold acceleration in throughput, enabling IFIs to process compliance queues that would otherwise span weeks within a single operational day. The clause-level retrieval precision of 91.4% and AAOIFI rule recall of 88.7% compare favourably with supervised document classification benchmarks reported in analogous legal NLP studies [11, 15], and are expected to improve

further with domain-specific fine-tuning of the underlying embedding model.

Critically, the 72.3% steady-state cache hit rate—achieved after an initial corpus warming period of approximately 500 queries—validates the DKC layer as a genuinely effective operational component rather than a theoretical construct. At this hit rate, the majority of compliance queries in a mature deployment are resolved without invoking the full vector retrieval and LLM inference pipeline, yielding commensurate reductions in API cost and infrastructure load.

6. Challenges, Security, and Ethical Considerations

6.1 Prompt Injection and Adversarial Manipulation

The deployment of LLM-based compliance systems in financial institutions introduces novel attack surfaces that have no counterpart in traditional rule-based RegTech. Chief among these is the threat of prompt injection: a class of adversarial input in which a malicious actor embeds instruction sequences within a submitted contract document, potentially overriding the system prompt and causing the LLM to produce fabricated compliance approvals [26]. In the context of Shariah compliance, a successful prompt injection attack could result in a demonstrably non-Shariah-compliant financial instrument being certified as permissible—a consequential fraud with both financial and religious dimensions.

SCAF addresses prompt injection through a multi-layered defence strategy. Input sanitization filters are applied to all submitted documents prior to embedding, detecting and neutralizing common injection patterns including role-switching commands, delimiter manipulation, and instruction overrides. The system prompt is delivered via a protected, non-user-accessible channel and is not included in the token window visible to document content. Furthermore, a secondary validation LLM—operating under a distinct system prompt oriented toward adversarial detection rather than compliance reasoning—evaluates each primary LLM output for internally inconsistent or implausibly permissive determinations. Any output flagged by the secondary model is escalated directly to scholar review regardless of the primary model's verdict.

6.2 The 'Black Box' Problem and Explainability Requirements

A persistent challenge in deploying neural AI systems within institutional governance frameworks is the opacity of generative model reasoning. Shariah compliance determinations carry legal and religious authority; they must be explicable, auditable, and contestable. An LLM that produces a compliance verdict without transparent reasoning trails is institutionally and doctrinally untenable, regardless of its statistical accuracy [27].

SCAF addresses this through mandatory structured output generation. Every compliance report produced by the system must include: (i) the specific AAOIFI standard and clause cited; (ii) the full text of the retrieved document chunk that supports the determination; (iii) a natural-language reasoning trace that maps the submitted clause to the identified standard; and (iv) a confidence score derived from the cosine similarity of the retrieved chunks to the query. These components collectively constitute an audit trail that a qualified Shariah scholar can review, validate, and annotate. The output format is designed to function as structured input to a scholar review workflow, not as a stand-alone compliance certificate.

6.3 Human-in-the-Loop Governance Model

The SCAF architecture explicitly rejects the model of fully autonomous Shariah compliance certification. Consistent with the jurisprudential principle that Ijtihad (independent legal reasoning) on novel financial instruments requires qualified scholarly authority [28], the framework mandates that all compliance determinations—whether positive (compliant) or negative (non-compliant with revision)—be reviewed and ratified by a certified member of the institution's SSB prior to contractual execution.

The Human-in-the-Loop (HitL) integration is operationalized through a structured scholar review interface that presents the SCAF output alongside its

full reasoning trace, retrieved evidence, and confidence metrics. The scholar may: (i) ratify the determination unchanged; (ii) modify the revision suggestion with annotation; or (iii) override the determination with a documented scholarly reasoning. All outcomes are logged in an immutable audit ledger, creating an institutional record that satisfies both regulatory reporting requirements and the retrospective review needs of SSB governance [29].

Under this model, the effective workload reduction accruing to scholars is substantial. In the simulated evaluation, 85.8% of SCAF determinations were ratified without modification, requiring only brief scholar review. The remaining 14.2%—cases involving novel instrument structures, ambiguous clause language, or jurisdictional variation in scholarly interpretation—were escalated for full SSB deliberation. This triaging function constitutes the primary operational value proposition of SCAF: not the replacement of scholarly authority, but the concentration of that scarce human resource on the genuinely novel, complex, and consequential cases that require it.

6.4 Jurisdictional and Madhab Heterogeneity

A non-trivial challenge in developing a universal Shariah compliance tool is the doctrinal plurality inherent in Islamic jurisprudence. The four major Sunni schools of law (Madhabs)—Hanafi, Maliki, Shafi'i, and Hanbali—differ in their positions on numerous financial contract elements, and their respective positions have been differentially adopted by SSBs across jurisdictions [30]. A compliance determination acceptable under the Hanbali-influenced rulings prevalent in Gulf Cooperation Council (GCC) states may conflict with the Shafi'i-influenced positions of Malaysian regulatory bodies.

SCAF addresses this through a jurisdiction-tagging mechanism applied at both the corpus and the query level. Each knowledge base document is tagged with its jurisdictional provenance and the Madhab tradition it reflects. At query time, the IFI's operational jurisdiction is specified as a query parameter, constraining retrieval to the relevant scholarly tradition. For instruments intended for multi-jurisdictional issuance, the system supports parallel retrieval across multiple jurisdictional indices, flagging points of inter-Madhab divergence for explicit SSB resolution.

7. Conclusion and Future Work

7.1 Summary of Contributions

This paper has presented SCAF, a RAG-LLM framework for the automation of first-pass Shariah compliance verification in Islamic banking and finance. The framework addresses a genuine and growing institutional bottleneck: the inability of the global community of qualified Shariah scholars to meet the compliance review demands of an industry growing at an annual rate of 10–15% [1]. Through the integration of a curated authoritative corpus, bilingual dense retrieval, LLM-grounded reasoning, Dynamic Knowledge Caching, and a structured Human-in-the-Loop governance model, SCAF offers a technically sound and institutionally viable pathway toward scalable Shariah governance.

The simulated evaluation demonstrates that SCAF can reduce average contract review time by approximately 240-fold, achieve clause-level retrieval precision exceeding 91%, and reduce computational operational costs by up to 45% through dynamic caching—while maintaining mandatory human scholarly oversight of all determinations. The case study evaluation of a Mudaraba contract confirms the system's capacity to detect substantive Shariah violations and generate compliant clause revisions grounded in authoritative AAOIFI evidence.

7.2 Limitations

Several limitations circumscribe the scope of the present work. First, the evaluation reported herein is simulation-based; empirical validation against real-world IFI contract portfolios, with ground-truth determinations provided by certified SSB scholars, represents a necessary and imminent next step. Second, the Arabic NLP capability of the underlying embedding

model, while demonstrably functional, has not yet achieved parity with English-language performance on Islamic legal texts—a gap that constrains the system's utility for corpus documents and contracts composed exclusively in classical Arabic. Third, the framework has not yet been evaluated for Shia Islamic finance instruments, which operate under distinct jurisprudential foundations; extension to this domain requires both corpus expansion and scholarly consultation.

7.3 Future Research Directions

Several productive avenues for future research emerge from this work:

- Domain-Specific Fine-Tuning: The development of a purpose-built Shariah Legal Language Model (SLLM) trained from the ground up on Arabic and English Islamic jurisprudence corpora, leveraging instruction-tuning techniques to align model outputs with authenticated scholarly reasoning patterns.
- Empirical Validation with IFI Partners: Deployment of SCAF in partnership with one or more live IFIs for prospective evaluation against real contract review workflows, enabling ground-truth accuracy measurement and SSB scholar feedback loops.
- Federated Knowledge Base Architecture: Design of a privacy-preserving federated knowledge base enabling IFIs to contribute proprietary fatwa corpora and historical compliance determinations without exposing confidential institutional data, thereby enriching the shared knowledge base while respecting competitive and regulatory constraints.
- Multi-Modal Extension: Integration of Arabic handwritten document recognition for jurisdictions where contracts are executed in handwritten form, extending SCAF's input modality beyond digitally typed text.
- Blockchain Audit Trail Integration: Anchoring SCAF audit logs to an immutable distributed ledger, providing regulatory bodies and external auditors with cryptographically verifiable records of all compliance determinations and scholar ratification events.

In summation, SCAF represents a principled and practically grounded contribution to the emerging field of AI-augmented Islamic finance governance. The framework is designed not to supplant the indispensable authority of the Shariah scholar, but to amplify its reach—enabling a finite community of qualified jurists to exercise meaningful oversight over a global financial system of ever-increasing scale and complexity. Future empirical work will be critical in translating these theoretical contributions into verified institutional practice.

References

[1] Islamic Financial Services Board (IFSB), "Islamic Financial Services Industry Stability Report," IFSB, Kuala Lumpur, 2023.
[2] M. U. Chapra, Towards a Just Monetary System. Leicester: The Islamic Foundation, 1985.
[3] Z. Hasan, "Shariah governance in Islamic banks: practices, practitioners and praxis," Global Economy Journal, vol. 11, no. 2, pp. 1–25, 2011.
[4] I. Warde, Islamic Finance in the Global Economy, 2nd ed. Edinburgh: Edinburgh University Press, 2010.
[5] T. Brown et al., "Language models are few-shot learners," in Proc. NeurIPS, 2020, pp. 1877–1901.
[6] Z. Ji et al., "Survey of hallucination in natural language generation," ACM Computing Surveys, vol. 55, no. 12, pp. 1–38, 2023.
[7] P. Lewis et al., "Retrieval-augmented generation for knowledge-intensive NLP tasks," in Proc. NeurIPS, 2020, pp. 9459–9474.
[8] D. Arner, J. Barberis, and R. Buckley, "The evolution of FinTech: a new post-crisis paradigm?," Georgetown Journal of International Law, vol. 47, no. 4, pp. 1271–1319, 2016.
[9] T. Philippon, "The FinTech opportunity," BIS Working Papers No. 655, Bank for International Settlements, 2017.
[10] R. Loughran and B. McDonald, "Textual analysis in accounting and finance: a survey," Journal of

Accounting Research, vol. 54, no. 4, pp. 1187–1230, 2016.

[11] I. Chalkidis et al., "LEGAL-BERT: the muppets straight out of law school," in Findings of EMNLP, 2020, pp. 2898–2904.

[12] D. Tsarapatsanis and N. Aletras, "On the ethical limits of natural language processing on legal text," in Proc. ACL Workshop on Ethics in NLP, 2021, pp. 1–10.

[13] V. Karpukhin et al., "Dense passage retrieval for open-domain question answering," in Proc. EMNLP, 2020, pp. 6769–6781.

[14] K. Guu et al., "Retrieval augmented language model pre-training," in Proc. ICML, 2020, pp. 3929–3938.

[15] Z. Shao et al., "Enhancing retrieval-augmented large language models with iterative retrieval-generation synergy," in Findings of EMNLP, 2023, pp. 9248–9274.

[16] A. Abozaid, "The evolution of Shariah governance framework in Islamic financial institutions," International Journal of Islamic and Middle Eastern Finance and Management, vol. 9, no. 4, pp. 486–500, 2016.

[17] M. Laldin and H. Furqani, "Developing Islamic finance in the framework of maqasid al-Shariah," International Journal of Islamic and Middle Eastern Finance and Management, vol. 6, no. 4, pp. 278–289, 2013.

[18] M. K. Hassan and M. K. Lewis, Handbook of Islamic Banking. Cheltenham: Edward Elgar Publishing, 2007.

[19] M. Billah, "Artificial intelligence in Islamic finance," Journal of Islamic Accounting and Business Research, vol. 11, no. 9, pp. 1941–1950, 2020.

[20] AAOIFI, Shariah Standards for Islamic Financial Institutions. Manama: Accounting and Auditing Organization for Islamic Financial Institutions, 2021.

[21] AAOIFI, Financial Accounting Standards. Manama: Accounting and Auditing Organization for Islamic Financial Institutions, 2022.

[22] R. Smith, "An overview of the Tesseract OCR engine," in Proc. 9th International Conference on Document Analysis and Recognition (ICDAR), 2007, pp. 629–633.

[23] L. Wang et al., "Text embeddings by weakly-supervised contrastive pre-training," arXiv preprint arXiv:2212.03533, 2022.

[24] Y. Malkov and D. Yashunin, "Efficient and robust approximate nearest neighbor search using HNSW graphs," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 4, pp. 824–836, 2020.

[25] M. Ayub, Understanding Islamic Finance. Chichester: John Wiley & Sons, 2007.

[26] K. Greshake et al., "Not what you've signed up for: compromising real-world LLM-integrated applications with indirect prompt injection," in Proc. IEEE Security and Privacy Workshops, 2023, pp. 79–90.

[27] F. Doshi-Velez and B. Kim, "Towards a rigorous science of interpretable machine learning," arXiv preprint arXiv:1702.08608, 2017.

[28] I. Kamali, Shari'ah Analysis of Issues in Islamic Finance. Kuala Lumpur: Ilmiah Publishers, 2008.

[29] Basel Committee on Banking Supervision, Principles for the Sound Management of Operational Risk. Basel: Bank for International Settlements, 2011.

[30] M. Usmani, An Introduction to Islamic Finance. The Hague: Kluwer Law International, 2002.

[31] Khaliq, A., Ali, A., Ajaz, S., & Paul, F. H. (2025). Comparative Analysis of FinBERT and DistilRoBERTa for NLP-Based Financial Insights in Pakistan's Stock Market. Spectrum of Engineering Sciences, 695-713.