



DFAS-AML: An Integrated Conceptual Governance Framework for Ethical, Explainable, and Sovereign-Sensitive AI in Anti-Money Laundering Systems

Hasan Mohamed Husain Alaali* 
DFAS Research Center, Bahrain

Received: July 19, 2026

Revised : August 7, 2026

Accepted : October 13, 2026

Online : October 14, 2026

Abstract

Artificial intelligence is increasingly embedded in anti-money laundering (AML) systems, yet improved analytical capability does not by itself ensure explainability, accountable human intervention, ethical resilience, jurisdiction-sensitive processing, or decision traceability. Existing research and regulatory instruments address these concerns but remain fragmented across technical, legal, organizational, and data-governance domains. This study adopts a normative conceptual and design-oriented methodology. It synthesizes peer-reviewed research and authoritative regulatory instruments, translates identified governance problems into operational requirements, derives corresponding governance mechanisms, and evaluates the resulting architecture through conceptual coherence, functional completeness, traceability, regulatory relevance, sovereign adaptability, operational interpretability, and illustrative stress scenarios. The analysis identifies five interdependent governance requirements: explanation governance, accountable human override, ethical resilience, sovereign calibration, and controlled decision traceability. These requirements are integrated into the Dynamic Financial Applied Meta-Science for Anti-Money Laundering (DFAS-AML) through five coordinated components: A-AMLEEF, A-AMLOCC, A-AMLD, A-SCEAML, and A-AMIL. The principal analytical finding is that effective governance of AI-enabled AML requires governance of the complete decision lifecycle rather than the detection model alone, linking explanation adequacy, ethical and jurisdictional assessment, authorized intervention, institutional action, and tamper-evident decision reconstruction. DFAS-AML provides an AML-specific governance architecture that integrates functions otherwise addressed across separate governance domains. It complements rather than replaces existing AML and AI-governance instruments and does not establish empirical effectiveness. The architecture remains conceptually evaluated and requires subsequent expert, institutional, empirical, and regulatory validation.

Keywords: *Anti-Money Laundering, AI Decision Governance, Explainability Governance, Accountable Human Override, Ethical Resilience, Sovereign Calibration, Decision Traceability.*

INTRODUCTION

AI-Enabled AML and the Governance Problem

Artificial intelligence (AI) is increasingly embedded in anti-money laundering (AML) operations through transaction monitoring, anomaly detection, alert prioritization, network analysis, and risk classification. The Financial Action Task Force (FATF) recognizes that advanced analytics and other new technologies can improve the speed, quality, and effectiveness of AML/CFT processes, while the Basel Committee on Banking Supervision (BCBS) identifies artificial intelligence and machine learning among technologies increasingly used across banking activities (FATF, 2021; BCBS, 2024). These developments extend AML capability beyond conventional rule-based monitoring toward systems capable of identifying complex relationships and changing patterns of suspicious activity.

The governance problem, however, is not equivalent to the detection problem. A technically successful model may generate a useful risk signal while leaving unresolved whether the output



can be adequately explained, who is authorized to challenge it, what evidence is sufficient for escalation, how potentially discriminatory or disproportionate effects should be addressed, and how decisions should be reconstructed for audit or regulatory review. Automated suspicious-transaction monitoring therefore creates legal and institutional questions alongside technical ones (Turksen et al., 2024). AI-supported AML processing may also affect privacy, data protection, non-discrimination, and procedural rights, particularly when opaque models influence consequential investigative or compliance actions (Bertrand et al., 2021).

This distinction becomes more important as model complexity increases. Explainable artificial intelligence research in finance has expanded considerably, but AML remains comparatively underrepresented within that literature (Weber et al., 2024). The relevant challenge is consequently shifting from whether AI can detect suspicious patterns toward how institutions can govern AI-supported AML decisions throughout their lifecycle. Effective governance must connect detection with explanation, authorized human judgment, ethical safeguards, jurisdictional constraints, documentation, and accountable institutional action.

Fragmentation of Existing Governance Responses

Important elements of AI governance are already addressed across several bodies of literature and regulation. Explainable AI research examines transparency and interpretability (Weber et al., 2024), while organizational AI governance focuses on the structures, processes, roles, and responsibilities through which organizations direct and control AI activities (Mäntymäki et al., 2022). Human-oversight and rights-oriented scholarship further addresses intervention, privacy, fairness, proportionality, and procedural protection.

Regulatory instruments provide complementary mechanisms. The National Institute of Standards and Technology (NIST) AI Risk Management Framework organizes AI risk management around Govern, Map, Measure, and Manage functions (NIST, 2023). Regulation (EU) 2024/1689 establishes requirements for specified high-risk AI systems concerning risk management, documentation, record-keeping, transparency, and human oversight (European Union, 2024). The Basel Committee's principles for operational resilience and related supervisory guidance provide an additional institutional perspective on risk ownership, operational controls, and accountability within banking organizations (BCBS, 2021, 2024). Data-governance scholarship additionally identifies overlapping individual, organizational, and state claims over data sovereignty (Ryan et al., 2024).

These mechanisms nevertheless originate from distinct analytical and regulatory streams. Explainability does not independently allocate decision authority, human oversight does not automatically ensure traceability, and neither function alone resolves jurisdictional conflicts. Existing organizational AI-governance frameworks provide important integrative foundations, particularly through their treatment of governance structures, processes, and responsibilities (Mäntymäki et al., 2022). The unresolved issue addressed here is more specific: how these functions can be connected within an AML decision sequence in which an inadequate explanation, ethical concern, jurisdictional restriction, or authorized override changes the subsequent institutional action and its documentary requirements. The relevant problem is therefore fragmentation across operational governance functions rather than an absence of governance principles or integrative scholarship.

Research Gap and Research Question

Within AML, this fragmentation remains consequential. Existing scholarship identifies implementation difficulties in AML regulation (Zavoli & King, 2021), governance challenges associated with automated suspicious-transaction monitoring (Turksen et al., 2024), and

fundamental-rights concerns surrounding AI-supported AML systems (Bertrand et al., 2021). FATF guidance addresses responsible technological adoption within risk-based AML/CFT systems (FATF, 2021, 2026).

The specific research gap concerns the limited articulation of an integrated AML-specific architecture connecting these governance functions across the decision lifecycle. Existing regulatory instruments and organizational governance frameworks provide relevant requirements and principles but do not, individually, establish the particular combination of explanation-triggered escalation, assigned intervention authority, ethical and jurisdictional decision conditions, and reconstructable institutional action proposed in this study. The gap is therefore one of AML-specific operational integration, not governance availability.

The proposed DFAS-AML architecture builds on the author's earlier conceptual work, including the DFAS Governance Convergence Doctrine (DFAS-GC), the DFAS AI Accountability Protocol (DFAS-AAP), the DFAS Universal PostObjective framework (DFAS-UP), and the DFAS Governance Deployment Roadmap (DFAS-GDR) (Alaali, 2025a, 2025b). Related work has also examined governance convergence in the Persian Gulf Cooperation Council (PGCC) and emerging-market contexts (Alaali, 2026). These earlier contributions establish broader conceptual foundations for governance convergence, accountability, ethical adaptation, and institutional deployment. The present study's incremental contribution is their AML-specific operational integration into a coordinated architecture that connects explanation adequacy, ethical resilience, authorized human override, sovereign calibration, and decision traceability across the AML decision lifecycle. It does not claim to originate these governance principles or to establish their empirical effectiveness.

The study addresses the following research question:

How can ethical resilience, explanation adequacy, human override authority, sovereign sensitivity, and decision traceability be integrated into a unified conceptual governance framework for AI-enabled AML systems?

Research Contribution

This study develops DFAS-AML as a sector-specific conceptual governance framework designed to integrate these functions within a single AI-enabled AML decision architecture. Its intellectual lineage derives from the Dynamic Financial Applied Meta-Science (DFAS) governance program. DFAS-GC establishes governance convergence as an embedded governance logic; DFAS-AAP develops role-based override and explainability-escalation mechanisms; and DFAS-UP provides the basis for dynamic ethical recalibration (Alaali, 2025a). DFAS-GDR subsequently translates governance principles into institutional deployment structures (Alaali, 2025b), while the PGCC and emerging-market governance model extends financial AI governance toward accountability, operational control, regulatory alignment, and jurisdiction-sensitive governance (Alaali, 2026).

DFAS-AML does not treat this lineage as independent evidence of the external governance gap. Instead, it uses these prior constructs as design antecedents and specializes them for AML. Consistent with design-science reasoning, the contribution lies in developing and positioning an integrated governance artifact in response to an identified problem domain rather than claiming that its conceptual construction establishes empirical effectiveness (Gregor & Hevner, 2013). The resulting framework integrates five functions: governance requirements, explanation escalation, human override authority, sovereign calibration, and decision traceability. Its contribution is therefore integrative and operational: it converts independently documented governance problems into an AML-specific architecture whose empirical effectiveness remains subject to subsequent validation.

LITERATURE REVIEW

AI-Enabled AML: From Detection to Governance

AI-enabled AML systems extend conventional monitoring by enabling institutions to process large and heterogeneous datasets, identify nonlinear patterns, prioritize alerts, and examine relationships that may be difficult to capture through fixed rules. FATF identifies new technologies as capable of improving AML/CFT effectiveness and efficiency, including risk assessment, monitoring, information processing, and supervisory capabilities (FATF, 2021). BCBS similarly documents expanding AI and machine-learning applications across banking and emphasizes that digitalization can generate new vulnerabilities or amplify existing operational, strategic, reputational, and governance risks (BCBS, 2024).

AML therefore presents a dual problem. The first concerns analytical performance: whether a system can identify suspicious behavior accurately and efficiently. The second concerns institutional governability: whether resulting outputs can be lawfully interpreted, challenged, escalated, documented, and translated into accountable action. Turksen et al. (2024) demonstrate that automated suspicious-transaction monitoring raises issues involving transparency, explainability, integrity, and legal accountability. Bertrand et al. (2021) further show that machine-learning-based transaction monitoring can intensify fundamental-rights concerns relating to data processing and proportionality.

These concerns operate within an AML environment already characterized by substantial implementation difficulties. Zavoli and King (2021) demonstrate empirically that translating AML obligations into practice can produce significant operational challenges even without introducing complex AI systems. AI therefore does not replace the institutional AML problem; it adds another decision layer whose technical outputs must be incorporated into existing legal, compliance, investigative, and supervisory structures. The relevant unit of governance is consequently not the algorithm alone but the decision process through which an algorithmic signal becomes institutional action.

Explainability and Explanation Fidelity

Explainability is particularly important in finance because model outputs may influence decisions requiring justification to internal reviewers, auditors, regulators, customers, or other affected stakeholders. Weber et al. (2024), in a systematic review of explainable artificial intelligence (XAI) applications in finance, identify substantial growth in explainability research while finding AML comparatively understudied. Their review also demonstrates the diversity of explainability approaches and the increasing use of post hoc methods alongside intrinsically interpretable models.

The existence of an explanation, however, should not be equated with the reliability of that explanation. Post hoc explanations can simplify complex model behavior and may not fully reproduce the causal or computational process underlying an output. The Financial Stability Institute identifies inaccuracy, instability, and potentially misleading explanations among significant limitations of current explainability techniques in financial applications (Perez-Cruz et al., 2025). This distinction is critical for AML because an explanation may influence whether an investigator accepts an alert, requests further evidence, escalates a case, or initiates a consequential compliance action.

Explanation governance must therefore consider explanation fidelity, not merely explanation availability. An institution requires mechanisms for determining whether an explanation is sufficiently reliable, stable, relevant, and comprehensible for the decision being made. NIST (2023) similarly treats explainability and interpretability as characteristics that must be understood within the context and risk profile of an AI system rather than as isolated technical

properties.

For AML governance, explanation adequacy should therefore be linked to institutional consequences. A sufficiently supported explanation may permit ordinary review; uncertainty or instability may require enhanced scrutiny; and materially inadequate explanations may justify escalation or restriction of automated reliance. Explainability becomes a governance mechanism only when its adequacy affects the subsequent decision path.

Human Oversight, Accountability, and Decision Authority

Human oversight is frequently proposed as a safeguard against inappropriate automated decision-making, but human presence alone does not establish meaningful control. Effective oversight requires competence, information, authority, and the practical ability to intervene. Research on meaningful human control emphasizes that the level of human involvement should respond to system autonomy and decision context rather than relying on a nominal human-in-the-loop arrangement ([Methnani et al., 2021](#)).

Organizational AI governance reinforces this institutional dimension. [Mäntymäki et al. \(2022\)](#) conceptualize AI governance through organizational structures, processes, and relational mechanisms, indicating that responsible governance depends on how decision rights and responsibilities are allocated rather than solely on technical system characteristics. [NIST \(2023\)](#) similarly places accountability structures, roles, policies, and organizational processes within the Govern function of AI risk management.

The European Union (EU) AI Act further formalizes human oversight for systems falling within its high-risk regime. Article 14 requires relevant systems to be designed so that they can be effectively overseen by natural persons and links oversight measures to the risks, autonomy, and context of use ([European Union, 2024](#)). Effective oversight therefore implies more than observation: reviewers must be capable of understanding relevant limitations and intervening where necessary.

AML adds a further requirement because intervention can itself produce material consequences. Governance must determine who may accept, reject, suspend, modify, or escalate an AI-supported output; what conditions activate that authority; and where responsibility resides after intervention. This requirement is particularly important in light of automation bias, whereby decision-makers may place excessive reliance on automated recommendations and insufficiently scrutinize contradictory information ([Parasuraman & Manzey, 2010](#)). Without clearly assigned authority, adequate information, and documented intervention procedures, human oversight may create an accountability illusion in which a person formally participates without exercising meaningful control. Decision authority, escalation rights, documented responsibility, and safeguards against automation bias are therefore integral to meaningful human oversight.

Ethical Resilience and Fundamental Rights

AI-supported AML monitoring operates at the intersection of legitimate financial-crime prevention and potentially intrusive processing of personal and transactional information. [Bertrand et al. \(2021\)](#) demonstrate that AI-based AML/CFT monitoring raises substantial questions concerning privacy, data protection, proportionality, and other fundamental rights. These concerns become more significant where opaque systems influence investigations, restrictions, enhanced due diligence, or suspicious-activity reporting.

The EU AI Act reflects a broader regulatory movement toward risk-sensitive protection of health, safety, and fundamental rights, including governance requirements for systems classified within its high-risk categories ([European Union, 2024](#)). Responsible AI governance consequently requires institutions to examine not only whether a system performs its assigned function but also

whether its operation generates unacceptable effects.

For AML, ethical resilience is proposed here as the institutional capacity to identify, reassess, and respond to material ethical or rights-related risks as operating conditions change. Changes in model performance, data composition, customer populations, legal requirements, or decision thresholds may alter the ethical profile of a previously assessed system. This interpretation is consistent with the NIST AI Risk Management Framework's emphasis on continuous risk identification, measurement, management, and monitoring throughout the AI lifecycle (NIST, 2023). Ethical governance therefore cannot be reduced to a pre-deployment checklist.

A resilient mechanism requires observable triggers and defined responses. Evidence of potentially discriminatory outcomes, disproportionate intervention, unexplained differential treatment, or material deterioration in explanation quality should activate review, recalibration, escalation, restriction, or suspension according to the seriousness of the risk. Such concerns are relevant to the fundamental-rights implications of AI-supported AML monitoring identified by Bertrand et al. (2021) and to the broader risk-management obligations articulated by NIST (2023). Within the proposed DFAS-AML architecture, ethical resilience is therefore conceptualized as an ongoing decision function that connects identified ethical concerns with proportionate institutional responses.

Sovereignty, Data Governance, and Jurisdictional Control

AML systems frequently operate across institutional and national boundaries. Customer information, transaction records, model infrastructure, cloud services, analytics providers, and investigative processes may be distributed across jurisdictions with different requirements governing privacy, localization, disclosure, retention, access, and supervisory authority. FATF itself recognizes that AML/CFT standards must be implemented through measures adapted to differing legal, administrative, operational, and financial circumstances (FATF, 2026).

Data sovereignty further complicates this environment because authority over data cannot always be reduced to a single actor. Ryan et al. (2024) distinguish among individual, organizational, and state claims to data sovereignty and demonstrate the governance tensions that can emerge when these claims overlap. Gikay (2024) likewise highlights differences between comprehensive and sector-oriented approaches to AI regulation and the difficulty of designing rules that remain proportionate and adaptable as technology evolves.

Regulation (EU) 2024/1689 adds another layer of jurisdiction-sensitive obligations for AI systems within its scope, while existing privacy, financial, and AML regimes continue to operate concurrently. Cross-border AML governance therefore requires more than technical interoperability.

Sovereign sensitivity in this context concerns the capacity of a governance architecture to identify applicable jurisdictional constraints and determine how they affect permissible data processing, access, transfer, escalation, documentation, and institutional action. This requirement follows from the overlapping sovereignty claims identified by Ryan et al. (2024) and the regulatory differences discussed by Gikay (2024). Within the proposed DFAS-AML architecture, sovereign calibration is consequently treated as a decision-governance function rather than a uniform technical restriction. Identical analytical outputs may require different institutional responses where applicable legal obligations, supervisory authorities, or data-governance conditions differ. The proposed mechanism structures the identification and consideration of these constraints without claiming to resolve conflicts of law or establish legal compliance independently.

Traceability, Documentation, and Model Governance

Accountability depends on the ability to reconstruct how an AI-supported decision developed. This requires more than preserving the final outcome. Relevant records may include the model version, input context, generated output, explanation, uncertainty assessment, reviewer identity, escalation event, override decision, jurisdictional adjustment, and final institutional action. Such documentation supports the distinction between producing an analytical output and establishing an accountable institutional decision, consistent with organizational AI governance's emphasis on defined responsibilities and governance processes (Mäntymäki et al., 2022).

NIST (2023) places documentation, transparency, accountability, and lifecycle risk management within its broader governance architecture. The EU AI Act provides a more formal example through record-keeping requirements for specified high-risk AI systems. Article 12 requires technical logging capabilities intended to support appropriate traceability, while Article 14 connects human oversight with the ability of authorized persons to understand and intervene in system operation (European Union, 2024). BCBS (2024) similarly emphasizes robust governance, risk management, data safeguards, supervisory capability, and operational resilience as financial institutions increase their reliance on digital technologies.

Traceability is particularly important in AML because decisions may be reviewed after the original alert, potentially by different institutional actors. A record showing only that an alert existed does not establish why it was accepted, rejected, escalated, or overridden. The documentation and accountability principles reflected in NIST (2023) and the organizational governance mechanisms discussed by Mäntymäki et al. (2022) support the need to preserve information about decision processes and assigned responsibilities. Within DFAS-AML, this requirement is extended conceptually to decision-level reconstruction connecting machine outputs, explanations, human interventions, and final institutional actions.

Traceability must also coexist with lawful record-management obligations. Audit integrity should not be interpreted as indefinite or legally unrestricted retention. FATF Recommendation 11 establishes record-keeping expectations for financial institutions, while applicable data-protection and jurisdictional requirements may impose additional conditions on retention, access, and disposal (FATF, 2026). A defensible architecture should therefore preserve tamper-evident decision history within applicable legal retention periods, access controls, correction procedures, and lawful disposal requirements. Model governance and record governance are consequently treated as interconnected components of accountable AI-supported AML, without assuming that technical immutability overrides legal obligations.

Literature Synthesis and Remaining Governance Gap

The literature provides substantial but functionally dispersed mechanisms for governing AI-enabled AML. FATF establishes risk-based AML/CFT and technology guidance; NIST provides lifecycle AI risk-management functions; the EU AI Act addresses risk, transparency, logging, and human oversight; and academic literature develops organizational governance, explainability, fundamental-rights, and data-sovereignty perspectives (Bertrand et al., 2021; European Union, 2024; FATF, 2021, 2026; Mäntymäki et al., 2022; NIST, 2023; Ryan et al., 2024; Weber et al., 2024). The remaining gap concerns how these mechanisms operate together when an AI-supported AML decision evolves from model output to institutional action. Existing approaches do not provide an AML-specific sequence linking explanation assessment, authorized intervention, ethical review, jurisdictional calibration, and decision reconstruction. DFAS-AML addresses this integration requirement by organizing these functions within a unified governance architecture.

RESEARCH METHOD

Research Design

This study adopts a normative conceptual and design-oriented research design supported by integrative literature synthesis and comparative document analysis. The purpose is not to estimate causal effects or test the empirical performance of an AI system, but to derive and organize governance requirements for AI-enabled anti-money laundering (AML) decision processes. Conceptual research is appropriate where the research objective involves integrating previously fragmented constructs, specifying relationships among them, and developing a new conceptual architecture (Jaakkola, 2020). Design-science reasoning further supports the construction of purposeful artifacts in response to identified organizational problems while requiring explicit articulation of the problem, design knowledge, and contribution (Gregor & Hevner, 2013; Hevner et al., 2004).

The methodological sequence consists of four linked stages: identification and selection of relevant evidence; analytical coding of documented governance problems; derivation of governance requirements from that evidence; and construction and conceptual evaluation of the DFAS-AML architecture. External academic and regulatory evidence establishes the problem domain and design requirements. DFAS antecedents establish intellectual lineage only. This separation prevents the proposed framework from functioning as evidence for the governance problems it is designed to address.

Evidence Identification and Selection

Evidence was identified through an integrative search designed to capture peer-reviewed scholarship and authoritative regulatory or institutional documents relevant to AI-enabled AML governance. Academic searches covered Scopus, Web of Science, Google Scholar, and publisher databases, supplemented by targeted searches of FATF, BIS/BCBS, NIST, and European Union repositories. The principal search period covered 2020–2026, with earlier sources retained where necessary to establish methodological or theoretical foundations.

Search terms targeted AI-enabled AML and the governance dimensions examined in the study, including explainability, human oversight, accountability, fundamental rights, data and jurisdictional governance, traceability, model governance, and institutional implementation. Boolean combinations were adapted to database functionality. The complete search structure is reported in Appendix A.

Sources were eligible when they addressed at least one governance dimension material to the study. Purely technical detection studies without substantive governance implications were excluded. Priority was given to peer-reviewed research, primary regulatory instruments, international standard-setting bodies, and recent sources relevant to current governance requirements.

The search was integrative and design-oriented rather than a systematic review intended to establish an exhaustive evidence universe or meta-analytic sample. Evidence was screened for relevance to the governance problem and its capacity to inform design requirements. Accordingly, the study does not report retrospective Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) counts or claim exhaustive literature coverage where a contemporaneous systematic-review screening log was not maintained.

Analytical Coding and Evidence Synthesis

Selected materials were analyzed through conceptual coding informed by established approaches to qualitative document analysis and integrative literature review (Bowen, 2009; Snyder, 2019). Coding did not treat frequency of occurrence as evidence of importance. Instead,

sources were examined for governance problems, institutional consequences, required controls, relationships among governance functions, and qualifications or limitations affecting proposed governance responses.

Five recurrent problem categories were retained because they were independently supported by the reviewed evidence and directly relevant to consequential AML decisions: explanation inadequacy; unclear human intervention authority; ethical or fundamental-rights degradation; jurisdictional and sovereign conflict; and inadequate decision reconstruction.

Potentially contradictory or disconfirming evidence was considered by examining whether reviewed sources challenged the necessity, feasibility, scope, or assumed benefits of a proposed governance requirement. Particular attention was given to limitations of post hoc explanations, the possibility that nominal human oversight may reproduce automation bias, tensions between intervention authority and operational efficiency, competing jurisdictional obligations, and conflicts between audit preservation and lawful data-retention requirements. Such evidence was treated as a qualification or boundary condition rather than automatically excluded from the synthesis.

The analysis distinguished between evidence that identified a governance problem and evidence that supported a particular institutional response. Where a source established a risk without demonstrating the effectiveness of a proposed control, it was used to justify the corresponding design requirement rather than to validate the resulting DFAS-AML mechanism. This distinction limited the risk of converting normative recommendations into unsupported effectiveness claims.

The synthesis therefore produced governance requirements rather than predetermined DFAS components. Component construction occurred only after the external evidence had been organized into identifiable governance problems.

Evidence-to-Design Derivation

The central derivation logic was: External Evidence → Governance Problem → Design Requirement → DFAS-AML Mechanism

Each proposed mechanism was required to respond to a governance requirement independently identifiable in external literature or regulatory evidence. Explanation limitations generated a requirement for graduated explanation governance; ambiguous human authority generated a requirement for accountable intervention and escalation; rights-related risks generated a requirement for dynamic ethical review; cross-jurisdictional constraints generated a requirement for sovereign calibration; and reconstruction deficiencies generated a requirement for controlled decision traceability.

The DFAS research program entered only at the design stage. DFAS-GC, DFAS-AAP, and DFAS-UP provide antecedent concepts concerning governance convergence, explanation escalation, override accountability, and dynamic ethical recalibration ([Alaali, 2025a](#)). DFAS-GDR provides antecedent deployment logic for translating governance principles into institutional processes ([Alaali, 2025b](#)). The PGCC and emerging-market model provides prior financial-governance logic concerning accountability, operational control, regulatory alignment, and jurisdiction-sensitive governance ([Alaali, 2026](#)).

These sources establish intellectual provenance rather than independent confirmation of the identified problems. DFAS-AML therefore constitutes an AML-specific design response constructed from externally established requirements and prior DFAS design knowledge. This distinction follows design-science reasoning that separates the problem environment, existing knowledge base, and newly constructed artifact ([Gregor & Hevner, 2013](#)).

Conceptual Evaluation

DFAS-AML was evaluated against six predefined conceptual criteria: conceptual coherence, functional completeness, traceability, regulatory relevance, sovereign adaptability, and operational interpretability. Conceptual coherence assesses whether the five components operate according to compatible assumptions and decision logic. Functional completeness assesses whether the architecture addresses each governance requirement derived from the evidence. Traceability examines whether governance actions can be reconstructed across automated and human stages. Regulatory relevance assesses consistency with material requirements reflected in FATF guidance, the NIST AI Risk Management Framework (AI RMF), the EU AI Act, and financial-sector governance instruments. Sovereign adaptability assesses whether the architecture can accommodate materially different jurisdictional constraints without abandoning its core governance functions. Operational interpretability assesses whether institutional actors can identify what action is required, by whom, and under which governance condition.

The evaluation follows design-science principles that distinguish the conceptual construction and justification of an artifact from subsequent demonstration of its effectiveness in operational settings (Gregor & Hevner, 2013; Hevner et al., 2004). It is logical and comparative rather than statistical. A criterion is considered conceptually satisfied when the architecture contains an identifiable mechanism corresponding to the criterion and no unresolved internal contradiction prevents that mechanism from operating as specified.

Illustrative stress scenarios were used to examine how the proposed components would interact under governance tensions, including inadequate explanations, disputed overrides, ethical concerns, jurisdictional restrictions, and conflicting documentation obligations. These scenarios served as conceptual consistency checks rather than simulated evidence of institutional performance.

Methodological Boundaries

The methodology does not constitute empirical validation of DFAS-AML. No institutional deployment, controlled experiment, expert Delphi assessment, regulatory sandbox, or performance comparison was conducted. The findings therefore establish conceptual derivation, internal coherence, and regulatory positioning rather than effectiveness. Claims concerning improved AML outcomes, reduced compliance failures, institutional usability, implementation costs, or superior governance performance require subsequent independent empirical evaluation.

FINDINGS AND DISCUSSION

The DFAS-AML Conceptual Governance Framework

Derived Governance Requirements

The synthesis identified five governance requirements arising when an AI-generated AML signal enters an institutional decision process.

Table 1. Derived Governance Requirements and DFAS-AML Responses

Evidence-derived problem	Governance requirement	DFAS-AML response
Explanation inadequacy	Explanation governance	A-AMLEEF
Unclear intervention authority	Accountable override	A-AMLOCC
Rights/fairness degradation	Ethical resilience	A-AMLD
Jurisdictional conflict	Sovereign calibration	A-SCEAML
Decision reconstruction failure	Traceability	A-AMIL

Explanation governance responds to limitations in the fidelity, stability, and decision suitability of AI explanations (Perez-Cruz et al., 2025; Weber et al., 2024). Accountable override addresses the need for identifiable human authority and effective intervention (European Union, 2024; Mäntymäki et al., 2022). Ethical resilience responds to privacy, proportionality, discrimination, and procedural-rights risks (Bertrand et al., 2021). Sovereign calibration addresses jurisdictional conflicts affecting data and decision authority (FATF, 2026; Ryan et al., 2024), while traceability enables reconstruction of automated and human contributions to consequential decisions (European Union, 2024; NIST, 2023).

The five categories represent a conceptual consolidation of related governance problems rather than an exhaustive classification. Explanation fidelity, interpretability, and explanation stability were grouped under explanation governance; reviewer competence, intervention authority, and escalation responsibility under accountable override; privacy, fairness, proportionality, and procedural-rights concerns under ethical resilience; data sovereignty, cross-border restrictions, and jurisdictional obligations under sovereign calibration; and documentation, logging, auditability, and decision reconstruction under traceability. Detection accuracy, operational efficiency, and implementation costs were not retained as independent governance categories because they concern technical performance or empirical implementation questions beyond the primary scope of the proposed decision-governance architecture. They remain relevant to subsequent evaluation rather than being treated as resolved by conceptual integration. Together, these requirements define the functional specification from which the DFAS-AML architecture is constructed.

DFAS-AML Integrated Architecture

DFAS-AML translates the five requirements into an interdependent governance sequence. A-AMLEEF assesses explanation adequacy and determines whether escalation is required; A-AMLD evaluates material ethical conditions; A-SCEAML identifies jurisdictional constraints affecting permissible processing or action; A-AMLOCC assigns authorized intervention when a governance trigger occurs; and A-AMIL preserves the resulting decision history.

The architecture extends earlier DFAS principles concerning governance convergence, accountability, ethical recalibration, institutional deployment, and sovereign-sensitive financial governance (Alaali, 2025a, 2025b, 2026), specializing them for the AI-enabled AML lifecycle. AI-supported AML signal → explanation assessment → ethical and jurisdictional assessment → authorized review or intervention → institutional action → controlled traceability.

The sequence is recursive rather than strictly linear. New evidence, explanation instability, ethical deterioration, or jurisdictional change can reactivate an earlier governance stage and require reassessment before the decision proceeds. The proposed architecture therefore structures interactions among governance functions rather than assuming that completion of one stage permanently satisfies its associated requirement.

A-AMLEEF: Explainability and Explanation Fidelity

The Alaali AML Explainability Escalation Framework (A-AMLEEF) structures the assessment of whether an AI-supported AML output possesses sufficient explanatory quality for institutional reliance. Its purpose is not to require maximum interpretability in every case but to align explanation requirements with decision consequence and uncertainty.

Financial XAI research demonstrates that explanation techniques differ substantially in method and purpose, while AML remains comparatively underrepresented in the field (Weber et al., 2024). Regulatory analysis further indicates that post hoc explanations may be inaccurate, unstable, or misleading, making explanation availability an inadequate governance threshold

(Perez-Cruz et al., 2025). A-AMLEEF therefore incorporates explanation fidelity alongside intelligibility.

The mechanism proposes three governance states. Level 1 applies where the material basis of the output is sufficiently interpretable, stable, and consistent with available evidence for ordinary review. Level 2 applies where explanation is partial, unstable, materially uncertain, or dependent on interpretive assumptions; enhanced human assessment is required before consequential action. Level 3 applies where the output is materially opaque, contradictory, unreliable, or incapable of supporting defensible institutional reasoning; reliance on the automated output is restricted and escalation becomes mandatory within the proposed framework.

The logic extends the explanation-escalation principle developed in DFAS-AAP (Alaali, 2025a) while adding AML-specific fidelity requirements. Consistent with NIST (2023), explanation adequacy is treated contextually: the required degree of explanation depends on risk, affected actors, institutional use, and the consequence of error. The three levels constitute proposed governance classifications rather than empirically validated thresholds.

A-AMLOCC: Human Override and Decision Authority

The Alaali AML Override Command Chain (A-AMLOCC) translates human oversight requirements into a proposed structure of explicit decision authority. Organizational AI governance requires structures and processes capable of assigning responsibility for AI-related decisions (Mäntymäki et al., 2022), while meaningful human control requires intervention capacity rather than passive observation (Methnani et al., 2021). Article 14 of the EU AI Act similarly connects effective human oversight with the ability of authorized persons to understand limitations and intervene where appropriate (European Union, 2024).

A-AMLOCC is designed to activate when predefined governance conditions are present. These include Level 2 or Level 3 explanation status, material inconsistency between model output and available evidence, significant ethical or rights concerns, jurisdictional conflict, detected model or data anomaly, or a decision whose consequences exceed the authority of the current reviewer. Authority follows a proposed role-based escalation chain rather than an unrestricted human override. The first authorized reviewer may confirm, request additional evidence, or escalate. Higher authority may suspend reliance on the automated output, require independent review, modify the permissible decision pathway, or authorize an alternative action. Decisions involving unresolved legal or sovereign conflict require escalation to the relevant compliance, legal, or jurisdictional authority.

Every intervention is expected to identify the trigger, authorized actor, evidence considered, action taken, and resulting responsibility. This design also addresses automation bias, which may lead decision-makers to over-rely on automated recommendations despite contradictory information (Parasuraman & Manzey, 2010). A-AMLOCC therefore distinguishes meaningful intervention from nominal human participation by requiring identifiable authority, evidentiary justification, and documented accountability. It extends the role-based override logic of DFAS-AAP (Alaali, 2025a) into an AML-specific command structure without claiming that the proposed controls have been empirically demonstrated to eliminate automation bias.

A-AMLD: Ethical Resilience

The Alaali AML Governance Doctrine (A-AMLD) proposes ethical resilience as a continuing governance function across the AML decision lifecycle. AI-supported transaction monitoring can affect privacy, proportionality, non-discrimination, and procedural rights, particularly when automated classifications contribute to consequential compliance action (Bertrand et al., 2021). The EU AI Act similarly embeds fundamental-rights protection within its risk-oriented governance

structure ([European Union, 2024](#)).

A-AML_D operationalizes ethical resilience through five stages: Trigger → Assessment → Action → Escalation → Documentation.

A trigger arises when evidence indicates a material ethical or rights-related concern. Relevant triggers include unexplained differential treatment, deterioration in model performance across relevant groups, disproportionate intervention, materially deficient explanation, inappropriate data use, or a significant change in operating context. Assessment determines the seriousness, affected interests, evidentiary strength, and reversibility of the potential harm.

The resulting action may involve enhanced review, recalibration, additional evidence requirements, restriction of automated reliance, or temporary suspension. Material or unresolved risks activate escalation through A-AML_{OCC}. The decision and supporting reasoning are subsequently recorded through A-AML.

This logic derives its dynamic ethical lineage from DFAS-UP, which conceptualizes ethical governance as subject to continuing contextual reassessment rather than static objective compliance ([Alaali, 2025a](#)). Within DFAS-AML, ethical resilience is bounded by observable triggers and institutional responses rather than treated as an unrestricted normative principle. The proposed sequence establishes conceptual response conditions but does not demonstrate that these interventions reduce discriminatory outcomes or other ethical harms in operational AML environments.

A-SCEAML: Sovereign Calibration

The Alaali Sovereign Calibration Engine for AML (A-SCEAML) addresses circumstances in which technically identical AML processes are subject to materially different jurisdictional conditions. Data sovereignty may involve overlapping claims by individuals, organizations, and states, creating governance conflicts over access, processing, control, and transfer ([Ryan et al., 2024](#)). Regulatory approaches to AI also differ substantially across jurisdictions in scope, institutional structure, and legal technique ([Gikay, 2024](#)).

FATF standards provide a common international AML/CFT baseline while allowing implementation within different legal, administrative, and operational environments ([FATF, 2026](#)). A-SCEAML therefore does not alter FATF obligations or determine applicable law. Its proposed function is to identify jurisdiction-sensitive constraints that affect how an AI-supported AML decision may be processed institutionally.

Calibration is triggered when a decision involves cross-border data, external infrastructure, conflicting retention requirements, restricted transfers, differing disclosure obligations, foreign supervisory access, or other jurisdictionally material conditions. The mechanism identifies the relevant governance constraint and structures adjustments to permissible processing, escalation, access, documentation, or transfer pathways. Unresolved conflicts are escalated rather than algorithmically resolved.

The mechanism extends the sovereign-sensitive financial governance logic developed in the PGCC and emerging-market model, which emphasizes that effective financial AI governance must remain responsive to jurisdictional and institutional differences ([Alaali, 2026](#)). A-SCEAML specializes this principle for AML by making sovereign calibration an explicit decision-stage function. Its conceptual specification does not establish independent legal authority or demonstrate compliance across jurisdictions.

A-AML: Traceability and Controlled Record Integrity

The Alaali AML Integrity Ledger (A-AML) proposes controlled reconstruction of the governance history surrounding an AI-supported AML decision. [NIST \(2023\)](#) emphasizes

documentation and accountability across the AI lifecycle, while the EU AI Act establishes record-keeping and logging requirements for specified high-risk systems (European Union, 2024). BCBS (2024) likewise emphasizes governance, data management, operational resilience, and effective oversight as digital technologies become more deeply embedded in banking.

A-AMIL is designed to record governance events rather than merely preserve the final AML outcome. Relevant entries include the model and version used, decision context, output, explanation status, identified uncertainty, ethical or jurisdictional trigger, reviewer identity and authority, escalation path, override action, supporting rationale, and final institutional disposition. Record integrity is tamper-evident rather than absolutely immutable. This distinction accommodates accountability alongside lawful correction, retention, access, and disposal requirements. Changes to protected records should themselves be traceable, while authorized lifecycle controls determine retention and deletion according to applicable obligations. FATF Recommendation 11 provides an AML-specific record-keeping reference point, while applicable jurisdictional requirements may impose additional conditions (FATF, 2026).

A-AMIL therefore combines evidentiary integrity with controlled information governance. Its institutional lineage derives partly from DFAS-GDR, which emphasizes auditable deployment and governance integration (Alaali, 2025b). Within DFAS-AML, traceability functions as the proposed connective layer through which explanation assessment, ethical review, sovereign calibration, escalation, and human intervention can subsequently be reconstructed and evaluated. The design specifies the information and accountability relationships required for reconstruction without claiming that a technical ledger has been implemented or its integrity independently tested.

Application of the six predefined criteria indicates conceptual coverage of the identified governance requirements, while empirical performance remains undetermined. Conceptual coherence is supported by differentiated but connected functions: A-AMLEEF assesses explanation adequacy, A-AMLD addresses ethical conditions, A-SCEAML evaluates jurisdictional constraints, A-AMLOCC governs authorized intervention, and A-AMIL preserves the decision record. Functional completeness is supported because each evidence-derived problem has a corresponding mechanism. This denotes coverage of the defined problem set rather than completeness across all possible AML governance risks. Traceability is provided through A-AMIL, linking model outputs, explanations, governance triggers, interventions, recalibrations, and final actions, consistent with broader documentation and record-keeping expectations (European Union, 2024; NIST, 2023). Regulatory relevance derives from correspondence with risk-based AML governance, AI risk management, human oversight, documentation, and institutional accountability requirements (BCBS, 2024; European Union, 2024; FATF, 2021, 2026; NIST, 2023). Sovereign adaptability is provided through jurisdiction-sensitive calibration under A-SCEAML, while operational interpretability derives from linking governance conditions to identifiable actors and actions. These findings establish conceptual and comparative plausibility, not operational effectiveness.

Comparative Positioning

DFAS-AML is positioned as a complementary, sector-specific conceptual governance architecture rather than an alternative to existing governance instruments. FATF establishes the international risk-based AML/CFT foundation through Recommendation 1 on risk assessment and the risk-based approach, Recommendation 10 on customer due diligence, Recommendation 11 on record-keeping, Recommendation 18 on internal controls, and Recommendation 20 on suspicious transaction reporting (FATF, 2026). FATF also addresses opportunities and governance considerations associated with responsible technological adoption (FATF, 2021). DFAS-AML does not replace these obligations. Instead, it proposes a decision-governance sequence connecting AI-

generated risk signals with explanation assessment, ethical and jurisdictional review, authorized intervention, and documented institutional action.

The NIST AI Risk Management Framework provides cross-sector lifecycle governance through Govern, Map, Measure, and Manage (NIST, 2023). Relevant subcategories include GOVERN 2.1 concerning documented roles and responsibilities, MAP 1.1 concerning the intended purpose and context of AI systems, MEASURE 2.8 concerning the examination of fairness and bias, and MANAGE 1.3 concerning responses to identified risks. DFAS-AML builds on these general governance functions by proposing AML-specific escalation conditions, explanation adequacy levels, intervention authority, sovereign calibration, and decision reconstruction. These mechanisms represent conceptual specialization rather than an alternative risk-management standard.

The EU AI Act establishes binding requirements for systems falling within its applicable provisions, including Article 9 on risk management, Article 12 on record-keeping, Article 13 on transparency, and Article 14 on human oversight (European Union, 2024). These provisions are relevant to the design of accountable AI governance, although their applicability to a particular AML system depends on its legal classification and circumstances. DFAS-AML proposes operational relationships among these governance concerns without asserting that all AML systems are automatically classified as high-risk under the Act.

Within banking governance, the Basel Committee's *Digitalisation of Finance* report examines the implications of emerging technologies for banks and supervisors, including governance, risk management, operational resilience, data governance, and supervisory capacity (BCBS, 2024). The Basel Committee's *Principles for Operational Resilience* further provide relevant institutional reference points, particularly Principle 1 on governance, Principle 2 on operational risk management, and Principle 3 on business continuity planning and testing (BCBS, 2021). These principles address institutional resilience rather than prescribing the particular AI-supported AML decision sequence proposed here. Related Bank for International Settlements (BIS) research examines the limitations and regulatory implications of AI explainability (Perez-Cruz et al., 2025), while organizational AI governance literature emphasizes structures, processes, and responsibility allocation (Mäntymäki et al., 2022).

DFAS-AML's distinguishing contribution is therefore AML-specific operational integration and sectoral specialization. It connects explanation adequacy, ethical resilience, human override authority, sovereign sensitivity, and decision traceability within a proposed decision lifecycle while preserving the independent authority and scope of existing regulatory instruments. Its conceptual architecture neither establishes regulatory compliance nor demonstrates superior institutional performance. Implementation remains subject to applicable legal, supervisory, and organizational requirements.

CONCLUSION

The increasing integration of artificial intelligence into anti-money laundering (AML) systems introduces governance challenges that extend beyond detection accuracy and analytical efficiency. This study addressed how ethical resilience, explanation adequacy, human override authority, sovereign sensitivity, and decision traceability can be integrated into a unified conceptual governance framework for AI-enabled AML systems.

The research question is addressed through five interdependent mechanisms within DFAS-AML. The Alaali AML Explainability Escalation Framework (A-AMLEEF) structures explanation adequacy assessment and risk-sensitive escalation. The Alaali AML Override Command Chain (A-AMLOCC) assigns intervention authority, establishes escalation responsibilities, and requires documented justification for human decisions. The Alaali AML Governance Doctrine (A-AMLGD)

structures continuing ethical assessment and proportionate responses to emerging rights-related concerns. The Alaali Sovereign Calibration Engine for AML (A-SCEAML) identifies jurisdictional constraints and their implications for permissible processing and institutional action. Finally, the Alaali AML Integrity Ledger (A-AMIL) proposes controlled, tamper-evident documentation enabling reconstruction of automated outputs, governance assessments, human interventions, and final institutional decisions.

These mechanisms are integrated through a common decision sequence connecting AI-generated AML signals with explanation assessment, ethical and jurisdictional review, authorized intervention, institutional action, and controlled traceability. The sequence permits reassessment when explanation reliability deteriorates, ethical risks emerge, jurisdictional conditions change, or additional human authority becomes necessary. The framework therefore proposes governance for the complete AI-supported AML decision lifecycle rather than treating its individual safeguards as independent controls.

The principal theoretical contribution is the specification of an AML-specific governance architecture that connects established but functionally dispersed governance requirements through explicit decision relationships, escalation conditions, and accountability mechanisms. Its contribution lies in operational integration and sectoral specialization rather than the introduction of new regulatory obligations or a new theory of financial-crime detection.

From a practical perspective, DFAS-AML provides a conceptual basis for defining institutional responsibilities, explanation-review requirements, override procedures, ethical escalation conditions, jurisdiction-sensitive controls, and decision-documentation expectations. It may inform governance design by compliance functions, model-risk teams, internal auditors, legal specialists, and supervisory authorities. However, these applications remain proposed uses rather than demonstrated institutional outcomes.

DFAS-AML complements existing AML regulation, AI risk-management frameworks, and banking-governance instruments without replacing their requirements or independently establishing regulatory compliance. Its evaluation establishes conceptual coherence, functional coverage, and comparative regulatory relevance, but not empirical effectiveness. Independent expert assessment, scenario-based testing, institutional implementation, and cross-jurisdictional evaluation remain necessary to determine its feasibility, operational performance, implementation costs, and potential unintended consequences.

Theoretical Contribution

DFAS-AML conceptualizes AI-enabled AML governance as an interdependent decision-governance sequence. Existing literature establishes distinct concerns involving financial explainability, organizational AI governance, fundamental rights, data sovereignty, and automated transaction monitoring ([Bertrand et al., 2021](#); [Mäntymäki et al., 2022](#); [Ryan et al., 2024](#); [Turksen et al., 2024](#); [Weber et al., 2024](#)). The theoretical contribution lies in specifying their interaction as an AI-generated AML signal progresses toward institutional action.

The architecture extends earlier DFAS concepts concerning governance convergence, accountability, dynamic ethical recalibration, institutional deployment, and sovereign-sensitive financial governance ([Alaali, 2025a, 2025b, 2026](#)). Within DFAS-AML, explanation inadequacy can activate escalation, ethical risk can alter permissible reliance, jurisdictional constraints can modify the decision pathway, human intervention is assigned explicit authority, and governance events remain reconstructable. DFAS-AML therefore advances a governance-integration proposition rather than a new detection theory or regulatory regime.

Practical and Regulatory Implications

DFAS-AML provides practitioners with a structured governance pathway after an AI-supported AML signal is generated. Compliance functions can use it to define escalation criteria and responsibility boundaries; model-risk functions can connect explanation instability and technical anomalies to governance responses; internal audit can reconstruct decision pathways; and legal or data-governance functions can address jurisdictional constraints.

For supervisors, the architecture shifts attention from model performance alone toward the governance surrounding consequential AI-supported decisions. This remains compatible with FATF's risk-based AML/CFT approach, NIST's broader AI risk-management structure, applicable EU AI Act requirements, and BCBS banking-governance principles (BCBS, 2024; European Union, 2024; FATF, 2026; NIST, 2023).

Operational implementation requires institution-specific roles, thresholds, documentation standards, technical interfaces, and escalation authority. DFAS-AML therefore provides a governance architecture rather than a universally fixed operating procedure.

Boundary Conditions and Unresolved Tensions

Several tensions remain unresolved at the conceptual stage. Explanation fidelity may be difficult to establish for complex models, and greater interpretability can conflict with performance, security, or operational constraints (Perez-Cruz et al., 2025; Weber et al., 2024). Human intervention can mitigate automation risk while introducing inconsistency, delay, expertise limitations, or new accountability ambiguity. Jurisdictional calibration may identify conflicts that an institution cannot resolve internally, particularly where sovereign or regulatory obligations are incompatible (Ryan et al., 2024). Traceability requirements may also conflict with lawful retention, correction, privacy, or deletion obligations.

Institutional feasibility represents an additional boundary. Governance mechanisms that are conceptually coherent may impose substantial staffing, technology, documentation, or coordination costs. DFAS-AML therefore cannot establish that integration will improve AML effectiveness or institutional efficiency without independent implementation evidence. These tensions define conditions for subsequent expert assessment, simulation, regulatory sandbox testing, and institutional validation.

LIMITATION & FURTHER RESEARCH

The principal limitation of this study follows directly from its conceptual and design-oriented methodology. DFAS-AML was derived from academic literature, regulatory and institutional documents, comparative analysis, and structured design reasoning. The study therefore supports claims concerning conceptual construction, functional integration, and comparative governance positioning, but it does not establish that the proposed architecture improves AML effectiveness, reduces compliance failures, strengthens institutional accountability, or produces superior outcomes relative to existing governance arrangements.

No independent expert evaluation was conducted. The proposed components, escalation logic, governance triggers, and allocation of decision authority have therefore not been assessed by AML practitioners, compliance officers, model-risk specialists, auditors, regulators, or legal experts. Similarly, DFAS-AML has not been implemented within a financial institution or tested against operational constraints such as legacy systems, staffing capacity, decision latency, governance costs, data-access restrictions, or conflicting institutional responsibilities.

Future validation should proceed incrementally. A structured expert assessment should first evaluate conceptual completeness, clarity, regulatory relevance, and institutional feasibility. Controlled scenario-based evaluation can subsequently examine whether different users apply A-

AMLEEF, A-AMLOCC, A-AMLD, A-SCEAML, and A-AMIL consistently when confronted with explanation failures, ethical concerns, jurisdictional conflicts, and override decisions. Comparative case studies can test whether the architecture remains applicable across different institutions and regulatory environments. Simulation and regulatory-sandbox testing can then examine decision pathways under controlled conditions. Pilot implementation within an institutional AML environment would provide the strongest subsequent evidence regarding usability, implementation burden, governance performance, and unintended consequences.

REFERENCES

- Alaali, H. M. H. (2025a). *DFAS doctrine compendium: Philosophical, strategic, and systemic frameworks for AI governance, crisis intervention, and algorithmic control*. SSRN. <https://doi.org/10.2139/ssrn.5367894>
- Alaali, H. M. H. (2025b). Governance deployment roadmap: Implementation frameworks and adoption protocols for the DFAS Governance Convergence Doctrine. *Journal of Islamic Banking, Economics and Policy*, 2(1), 43–65. <https://doi.org/10.63740/bjffnd45>
- Alaali, H. M. H. (2026). AI governance in financial systems: A PGCC and emerging markets convergence framework for accountability, sovereignty, and operational control. *Journal of Countries Studies*, 1–58. <https://doi.org/10.22059/JCOUNTST.2026.412773.1476>
- Basel Committee on Banking Supervision. (2021). *Principles for operational resilience*. Bank for International Settlements. <https://www.bis.org/bcb/publ/d516.htm>
- Basel Committee on Banking Supervision. (2024). *Digitalisation of finance*. Bank for International Settlements.
- Bertrand, A., Maxwell, W., & Vamparys, X. (2021). Do AI-based anti-money laundering (AML) systems violate European fundamental rights? *International Data Privacy Law*, 11(3), 276–293. <https://doi.org/10.1093/idpl/ipab010>
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40. <https://doi.org/10.3316/QRJ0902027>
- Cinà, A. E., Grosse, K., Demontis, A., Vascon, S., Zellinger, W., Moser, B. A., Oprea, A., Biggio, B., Pelillo, M., & Roli, F. (2023). Wild patterns reloaded: A survey of machine learning security against training data poisoning. *ACM Computing Surveys*, 55(13s), Article 294. <https://doi.org/10.1145/3585385>
- European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Official Journal of the European Union. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- Financial Action Task Force. (2021). *Opportunities and challenges of new technologies for AML/CFT*.
- Financial Action Task Force. (2026). *The FATF recommendations: International standards on combating money laundering and the financing of terrorism & proliferation*.
- Gikay, A. A. (2024). Risks, innovation, and adaptability in the UK's incrementalism versus the European Union's comprehensive artificial intelligence regulation. *International Journal of Law and Information Technology*, 32, Article eaae013. <https://doi.org/10.1093/ijlit/eaae013>
- Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly*, 37(2), 337–355. <https://doi.org/10.25300/MISQ/2013/37.2.01>
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75–105. <https://doi.org/10.2307/25148625>
- Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review*, 10, 18–26. <https://doi.org/10.1007/s13162-020-00161-0>

- Mäntymäki, M., Minkkinen, M., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. *AI and Ethics*, 2, 603–609. <https://doi.org/10.1007/s43681-022-00143-x>
- Methnani, L., Aler Tubella, A., Dignum, V., & Theodorou, A. (2021). Let me take over: Variable autonomy for meaningful human control. *Frontiers in Artificial Intelligence*, 4, Article 737072. <https://doi.org/10.3389/frai.2021.737072>
- National Institute of Standards and Technology. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)* (NIST AI 100-1). U.S. Department of Commerce. <https://doi.org/10.6028/NIST.AI.100-1>
- Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, 52(3), 381–410. <https://doi.org/10.1177/0018720810376055>
- Perez-Cruz, F., Prenio, J., Restoy, F., & Yong, J. (2025). *Managing explanations: How regulators can address AI explainability* (FSI Occasional Papers No. 24). Financial Stability Institute, Bank for International Settlements.
- Ryan, M., Gürtler, P., & Bogucki, A. (2024). Will the real data sovereign please stand up? An EU policy response to sovereignty in data spaces. *International Journal of Law and Information Technology*, 32(1), Article eaae006. <https://doi.org/10.1093/ijlit/eaae006>
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Turksen, U., Benson, V., & Adamyk, B. (2024). Legal implications of automated suspicious transaction monitoring: Enhancing integrity of AI. *Journal of Banking Regulation*, 25, 359–377. <https://doi.org/10.1057/s41261-024-00233-2>
- Weber, P., Carl, K. V., & Hinz, O. (2024). Applications of explainable artificial intelligence in finance: A systematic review of finance, information systems, and computer science literature. *Management Review Quarterly*, 74, 867–907. <https://doi.org/10.1007/s11301-023-00320-0>
- Zavoli, I., & King, C. (2021). The challenges of implementing anti-money laundering regulation: An empirical analysis. *The Modern Law Review*, 84(4), 740–771. <https://doi.org/10.1111/1468-2230.12628>

Appendix A — Literature Search and Screening Protocol

A.1 Search Scope

The literature search covered academic and institutional evidence relevant to AI-enabled anti-money laundering governance. The principal publication period was 2020–2026, with earlier publications retained where required for methodological or theoretical foundations.

Academic searches covered Scopus, Web of Science, Google Scholar, IEEE Xplore, SpringerLink, ScienceDirect, Oxford Academic, and relevant publisher databases. Regulatory and institutional searches covered official repositories maintained by FATF, BIS/BCBS, NIST, and the European Union.

A.2 Search Structure

The search targeted combinations of terms relating to artificial intelligence, machine learning, anti-money laundering, suspicious transaction monitoring, explainability, human oversight, accountability, AI governance, fundamental rights, data sovereignty, cross-border data governance, model governance, traceability, and documentation.

Search combinations were constructed around the principal governance dimensions examined in the study and adapted to the functionality and indexing conventions of individual databases. Searches of regulatory and institutional repositories additionally targeted AI governance, AML/CFT technology, banking supervision, model governance, explainability, human oversight, documentation, and jurisdictional requirements.

A.3 Inclusion and Exclusion Criteria

Criterion	Included	Excluded
Domain	AI-enabled AML; consequential financial AI governance; banking AI governance	AI applications unrelated to financial or governance implications
Governance relevance	Explainability, human oversight, accountability, rights, sovereignty, traceability, model governance	Pure prediction or detection studies without material governance relevance
Source quality	Peer-reviewed journal articles, systematic reviews, authoritative regulatory and institutional documents	Promotional material, unsupported commentary, commercial claims
Time	Primarily 2020–2026	Earlier literature unless theoretically or methodologically necessary
Language	English	Sources without an accessible English version
Regulatory evidence	FATF, BCBS/BIS, NIST, EU, and comparable authoritative bodies	Unverified secondary descriptions of regulatory requirements
Duplication	Strongest and most directly relevant source retained	Duplicative sources adding no material analytical value

A.4 Screening Procedure

Screening proceeded through four stages:

Identification → Relevance Screening → Full-Text Eligibility Assessment → Analytical Inclusion

Title and abstract screening removed sources outside the defined governance scope. Full-text assessment determined whether a source supported an identifiable governance problem, design requirement, comparative criterion, or methodological procedure. Sources were not retained merely because they discussed AI or AML. Contradictory or limiting evidence was retained where it qualified explanation reliability, human oversight, jurisdictional compatibility, institutional feasibility, or record-governance assumptions.

The study used a targeted integrative search rather than a systematic review designed to estimate a complete literature population. PRISMA-style numerical screening counts were therefore not retrospectively constructed where no contemporaneous screening log existed. The final analytical corpus comprises sources satisfying the stated criteria and used in the manuscript's evidence synthesis and supporting matrices.

Appendix B — Evidence-to-Design Traceability Matrix

Source	External evidence	Governance problem	Derived requirement	DFAS-AML component	Governance rule
Weber et al. (2024)	XAI techniques differ in purpose and capability; AML remains comparatively underrepresented	Explanation availability does not ensure decision adequacy	Context-sensitive explanation governance	A-AMLEEF	Explanation quality determines review intensity
Perez-Cruz et al. (2025)	Post-hoc explanations may be inaccurate, unstable, or misleading	Institutional reliance on unreliable explanations	Explanation fidelity assessment	A-AMLEEF	Material uncertainty or instability prevents routine reliance
Turksen et al. (2024)	Automated suspicious-transaction monitoring creates transparency, integrity, accountability challenges	Automated outputs may lack defensible institutional governance	Review and escalation mechanisms	A-AMLEEF / A-AMLOCC	Consequential or inadequately explained outputs require authorized review
Mäntymäki et al. (2022)	Organizational governance requires structures, processes, and responsibilities	AI Human presence does not establish decision authority	Explicit role and responsibility allocation	A-AMLOCC	Intervention authority must be assigned before consequential action

Source	External evidence	Governance problem	Derived requirement	DFAS-AML component	Governance rule
Methnani et al. (2021)	Meaningful human control requires effective intervention capacity	Symbolic human-in-the-loop arrangements	Actionable human intervention	A-AMLOCC	Authorized actors must be capable of altering or suspending reliance
European Union (2024)	Relevant high-risk systems require human oversight, documentation, logging, and risk controls	Weak oversight and incomplete decision reconstruction	Oversight and traceable intervention	A-AMLOCC / A-AMIL	Material intervention must be attributable and reconstructable
Bertrand et al. (2021)	AI-based AML may implicate privacy, proportionality, non-discrimination, and procedural rights	Technically valid outputs may create unacceptable rights consequences	Continuing ethical assessment	A-AMLD	Material ethical or rights risk activates reassessment and possible escalation
Ryan et al. (2024)	Data sovereignty involves overlapping individual, organizational, and state claims	Cross-border processing may generate conflicting authority	Jurisdiction-sensitive governance	A-SCEAML	Processing conditions must be calibrated to applicable jurisdictional constraints
Gikay (2024)	AI regulatory approaches differ materially across jurisdictions	Uniform governance assumptions may be inappropriate	Regulatory adaptability	A-SCEAML	Governance pathways must accommodate materially different regulatory conditions
FATF (2021, 2026)	Technology use must remain consistent with risk-based AML/CFT obligations and differing national conditions	Technical capability does not displace AML accountability	AML-specific risk governance	A-AMLD / A-SCEAML	Technology remains subordinate to applicable AML/CFT obligations
NIST (2023)	AI risk governance requires lifecycle accountability, documentation, monitoring, and management	Fragmented lifecycle accountability	Continuous governance and traceability	A-AMLD / A-AMIL	Governance events must remain attributable across the lifecycle

Source	External evidence	Governance problem	Derived requirement	DFAS-AML component	Governance rule
BCBS (2024)	Financial digitalization creates governance, third-party, and operational risks	Outsourcing and technological data, complexity fragment accountability	Institutional and third-party governance	A-AMLD / A-AMIL	Institutional accountability persists despite external technology provision

Appendix C — DFAS-AML Decision Rules

C.1 A-AMLEEF Explanation Assessment

A-AMLEEF evaluates five dimensions:

1. Source Transparency
2. Decision-Factor Visibility
3. Contextual Consistency
4. Model Reliability
5. User Interpretability

Level	Decision condition	Required action
Level 1	Material factors are identifiable, sufficiently stable, contextually consistent, and understandable to the authorized reviewer	Routine human review; disposition recorded
Level 2	Explanation is incomplete, uncertain, unstable, partially inconsistent, or requires specialist interpretation	Additional evidence; enhanced compliance or model-risk review; consequential reliance deferred where necessary
Level 3	Output is materially opaque, contradictory, anomalous, unreliable, or unverifiable	Automated reliance suspended; senior compliance or technical intervention required

Level 1 does not authorize automatic acceptance of the underlying AML output. Level 3 does not establish that the transaction is legitimate or illegitimate; it establishes that the AI output lacks a sufficient basis for institutional reliance.

C.2 A-AMLOCC Authority Rules

Authority level	Primary authority	Permissible action
Operational Review	AML investigator or analyst	Assess, query, seek evidence, refer or escalate
Compliance Override	Senior AML or compliance officer	Confirm, reject, suspend, modify, or escalate proposed case action
Technical Intervention	Model-risk, data-governance, validation, or technical authority	Investigate anomalies, restrict model use, initiate revalidation, or modify approved technical conditions
Executive Regulatory	or Senior governance, legal, risk, or competent authority	Resolve high-impact, systemic, legal, regulatory, or sovereign conflicts

Escalation

A-AMLOCC is activated by one or more of the following conditions:

- A-AMLEEF Level 2 or Level 3 classification;
- material conflict between model output and verified evidence;
- suspected model drift, data poisoning, manipulation, or technical failure;
- material ethical or fundamental-rights concern;
- consequential customer impact;
- jurisdictional conflict;
- repeated or anomalous override patterns;
- uncertainty concerning decision ownership or authority.

Every intervention records the trigger, actor, authority level, evidence, action, rationale, approval where required, timestamp, and disposition.

C.3 A-AMLD Ethical Resilience Rules

Trigger → Assessment → Action → Escalation → Documentation

Ethical review is activated by evidence of disproportionate intervention, unexplained differential treatment, material deterioration in relevant group outcomes, deficient explanation affecting consequential action, inappropriate data use, or material contextual change.

The response may include enhanced review, additional evidence, recalibration, restriction of automated reliance, temporary suspension, or escalation through A-AMLOCC.

C.4 A-SCEAML Sovereign Calibration Rules

A-SCEAML assesses six governance dimensions:

Processing → Storage → Access → Explanation → Escalation → Reporting and Transfer

State	Condition	Response
Standard Processing	Applicable jurisdictional requirements are identified and satisfied	Processing continues under approved controls
Restricted Processing	Processing remains permissible subject to additional limitations	Apply access, transfer, disclosure, storage, or processing restrictions
Jurisdictional Escalation	Material legal or regulatory uncertainty exists	Refer to authorized legal, compliance, data-protection, or regulatory authority
Processing Suspension	Required authorization is absent or material incompatibility remains unresolved	Suspend affected processing or transfer pending resolution

A-SCEAML does not autonomously determine contested law.

C.5 A-AMIL Integrity Rules

A-AMIL records:

- case and model identifiers;
- model version and approval status;
- material data provenance;
- output and explanation level;
- reviewer identity and authority;
- escalation and override events;
- jurisdictional determinations;
- final disposition;
- corrections and reopening;
- retention and disposal events;
- audit activity.

Records are tamper-evident, not permanently immutable. Authorized amendment, correction, retention, and lawful disposal remain permissible when the action itself is documented and attributable.

Appendix D — Stress and Failure Scenarios

D.1 Synthetic Identity and Suspected Data Poisoning

Following a model update, a financial institution observes inconsistent alerts across connected accounts displaying similar identity and transaction characteristics. The anomaly may result from synthetic identities, model drift, data poisoning, deficient validation, or legitimate changes in behaviour. Training-data manipulation can selectively degrade machine-learning performance without producing complete system failure ([Cinà et al., 2023](#)).

A-AMLEEF assigns Level 2 where explanation reliability is uncertain and Level 3 where the output becomes materially inconsistent or unverifiable. A-AMLOCC activates technical and compliance review. Reviewers examine model version, training-data changes, feature distributions, validation evidence, provenance, and affected alerts. Where reliability cannot be established, authorized personnel may suspend the model update, restore an approved version, initiate revalidation, and reassess affected cases. A-AMIL preserves the complete governance sequence.

Failure condition: the architecture fails operationally if personnel formally escalate the alert but lack access to training provenance, model history, or authority to suspend the model.

Conflict condition: restoration of an earlier model may reduce the suspected manipulation risk while reintroducing known performance deficiencies. DFAS-AML does not determine automatically which model should prevail; the conflict requires documented technical and compliance judgment.

D.2 Cross-Border Data Governance Conflict

A multinational institution proposes transferring customer information to a foreign processing environment for AML analytics. The originating jurisdiction restricts particular transfers, the service provider permits infrastructure access from multiple jurisdictions, and the receiving jurisdiction imposes different retention or disclosure requirements.

A-SCEAML identifies the data categories, processing purpose, jurisdictions, provider-control structure, access rights, retention requirements, and competent authorities. Where compatibility cannot be established, the process enters Jurisdictional Escalation or Processing Suspension. Alternative arrangements may include minimization, pseudonymization, restricted access, local processing, federated analytics, or another authorized transfer mechanism. A-AMIL records the governing restrictions and determination.

Failure condition: A-SCEAML cannot produce a lawful operational pathway where applicable requirements are genuinely incompatible.

Conflict condition: AML effectiveness may favour broader information sharing while privacy, localization, or secrecy requirements restrict the same transfer. DFAS-AML identifies and escalates the conflict; it does not eliminate the underlying legal incompatibility.

D.3 High-Confidence Output with Misleading Explanation

An AML model produces a high-confidence suspicious-activity classification accompanied by a clear explanation. Subsequent review reveals that the explanation is unstable: small changes in input conditions produce materially different explanatory factors while the prediction remains similar.

A-AMLEEF does not classify the output as Level 1 merely because the explanation is understandable. Explanation fidelity and stability are material to institutional reliance. The case moves to Level 2 or Level 3 depending on the severity of the instability, activating model-risk review through A-AMLOCC ([Perez-Cruz et al., 2025](#)).

Failure condition: a reviewer may mistake comprehensibility for fidelity and approve the output despite an unreliable explanation.

Conflict condition: delaying action for explanation review may increase operational or financial-crime risk. The architecture requires documented proportionality between explanation uncertainty and the consequences of immediate or delayed action.

D.4 Human Override Conflict

An experienced AML investigator rejects a model-generated alert based on contextual knowledge. A second reviewer considers the automated evidence sufficiently strong to warrant escalation. Both positions are defensible from the available evidence.

A-AMLOCC requires the disagreement to move through the defined authority chain rather than allowing either the algorithm or an individual reviewer to prevail automatically. The original model output, first override, competing assessment, evidence, authority, and final decision are preserved through A-AMIL.

Failure condition: human oversight becomes symbolic if the reviewer cannot challenge the model, but it becomes uncontrolled discretion if overrides require no evidence or accountability.

Conflict condition: meaningful human authority and consistency of institutional decisions can pull in opposite directions. DFAS-AML manages this tension through attributable escalation rather than assuming that human judgment is inherently superior.

D.5 Traceability–Deletion Conflict

A closed AML case contains model outputs, explanations, reviewer comments, and jurisdictional determinations required for subsequent audit reconstruction. A lawful retention rule later requires disposal of specified records.

A-AMIL does not treat integrity as permanent preservation. Authorized disposal occurs according to the applicable retention regime, while the disposal event, authority, timing, and permissible residual metadata are recorded.

Failure condition: absolute immutability would create conflict with lawful deletion or retention obligations.

Conflict condition: excessive deletion can undermine accountability, while excessive preservation can violate data-governance requirements. The architecture therefore requires controlled lifecycle integrity rather than indefinite retention.

Appendix E — Detailed Regulatory Crosswalk

Governance dimension	FATF	NIST AI RMF	EU AI Act	BCBS/BIS	DFAS-AML
Primary scope	AML/CFT standards and risk-based guidance	Cross-sector AI risk management	Risk-based regulation	Banking AI governance, supervision and digitalization	AI-enabled AML decision governance
Risk governance	Risk-based AML/CFT approach	Govern, Map, Measure, Manage	Risk-management obligations for systems within scope	Prudential and operational governance	A-AMLD coordinates and AML-specific governance conditions
Explainability	Supports appropriate understanding and responsible use of new technologies	Addresses explainability, interpretability and transparency	Transparency and oversight requirements where applicable	Addresses explainability and model-governance concerns	A-AMLEEF links explanation fidelity to review and escalation
Human oversight	Human judgment remains material to risk-based AML processes	Accountability and defined organizational roles	Human oversight requirements for relevant high-risk systems	Institutional control and governance responsibilities	A-AMLOCC assigns intervention authority and escalation
Fundamental rights / ethical risk	Risk-based controls operate within applicable legal systems	Trustworthiness characteristics include fairness and accountability	Fundamental-rights protection is embedded in the regulatory architecture	Addressed through broader governance and conduct environment	A-AMLD provides continuing ethical-risk triggers and responses
Sovereignty / jurisdiction	Standards implemented through different national legal and institutional systems	Context considered through mapping and governance	Applies within defined jurisdictional scope	Recognizes EU fragmentation and cross-border financial risks	A-SCEAML calibrates processing to jurisdiction-specific constraints
Data	Relevant	to Data quality and	Data-governance	Data, third-party	A-SCEAML and

Governance dimension	FATF	NIST AI RMF	EU AI Act	BCBS/BIS	DFAS-AML
governance	effective and lawful AML/CFT technology use	governance addressed across lifecycle management	requirements apply to relevant risk systems	and infrastructure governance emphasized	A-AMIL connect data authority to decision governance
Traceability	AML record-keeping obligations	Documentation and lifecycle accountability	Logging and record-keeping requirements where applicable	Controls, documentation and supervisory review	A-AMIL connects model invocation, explanation, intervention, calibration and outcome
Intervention logic	Does not prescribe a DFAS-AML-style command chain	Requires accountability but remains cross-sectoral	Establishes oversight capabilities where applicable	Emphasizes effective institutional control	A-AMLOCC defines four functional authority levels
Explanation escalation	No dedicated three-level AML explanation architecture	No AML-specific three-level escalation structure	No DFAS-AML-specific explanation classification	No DFAS-AML-specific classification	A-AMLEEF Levels 1–3
Jurisdictional response states	National implementation accommodates jurisdictional variation	Contextual assessment	risk Legal obligations defined by regulatory scope	Cross-border risks recognized	A-SCEAML: Standard, Restricted, Escalation, Suspension
Decision-history integration	AML records	Lifecycle documentation	Relevant logging requirements	Governance and supervisory records	A-AMIL integrates machine and human governance events
Legal status	International AML/CFT standard-setting instrument	Voluntary management framework	risk-Binding regulation within applicable scope	International banking supervisory standards and guidance	Conceptual governance doctrine; not a legal substitute

DFAS-AML is complementary to these instruments. Regulatory correspondence does not establish legal equivalence, certification, automatic compliance, or superiority (FATF, 2021, 2026; NIST, 2023; European Union, 2024; BCBS, 2024).

Appendix A.1 – (DFAS-FEP) Authorship Classification Table for This Manuscript

Manuscript Component	DFAS-FEP Class	Justification
DFAS-AML Integrated Governance Doctrine	Class 1	The conceptualization of DFAS-AML as an integrated governance doctrine connecting explanation governance, accountable human override, ethical resilience, sovereign calibration, and controlled decision traceability across the AI-enabled AML decision lifecycle was independently developed by the author.

Evidence-to-Design Governance Architecture	Class 1	The architecture translating External Evidence → Governance Problem → Design Requirement → DFAS-AML Mechanism was independently structured by the author to separate externally established governance problems from the subsequent DFAS-AML design response.
Five-Component DFAS-AML Architecture	Class 1	The integration of A-AMLEEF, A-AMLOCC, A-AMLD, A-SCEAML, and A-AMIL as five interdependent AML governance mechanisms was independently developed by the author.
A-AMLEEF – Alaali AML Explainability Escalation Framework	Class 1	The AML-specific architecture governing explanation adequacy through explanation fidelity, decision consequence, uncertainty, and graduated escalation states was independently developed by the author.
A-AMLOCC – Alaali AML Override Command Chain	Class 1	The role-based governance architecture assigning intervention authority, escalation rights, responsibility, and documented accountability for human intervention in AI-supported AML decisions was independently developed by the author.
A-AMLD – Alaali AML Governance Doctrine	Class 1	The continuing ethical-resilience mechanism structured around Trigger → Assessment → Action → Escalation → Documentation was independently developed by the author for AI-enabled AML governance.
A-SCEAML – Alaali Sovereign Calibration Engine for AML	Class 1	The mechanism for identifying jurisdiction-sensitive constraints and calibrating permissible processing, escalation, access, documentation, and transfer pathways was independently developed by the author.
A-AMIL – Alaali AML Integrity Ledger	Class 1	The controlled traceability architecture for reconstructing automated and human governance events while distinguishing tamper-evident integrity from absolute immutability was independently developed by the author.
Integrated AML Decision-Governance Sequence	Class 1	The sequence AI-supported AML signal → explanation assessment → ethical and jurisdictional assessment → authorized review or intervention → institutional action → controlled traceability, including its recursive operation, was independently structured by the author.
Five Derived Governance Requirements	Class 1	The integration of Explanation Governance, Accountable Override, Ethical Resilience, Sovereign Calibration, and Traceability into a unified AML-specific governance specification was independently structured by the author from the identified evidence base.
Conceptual Evaluation Architecture	Class 1	The evaluation of DFAS-AML through conceptual coherence, functional completeness, traceability, regulatory relevance, sovereign adaptability, and operational interpretability was structured by the author for this manuscript.
AML-Specific Extension of Prior DFAS Governance Logic	Class 1	The specialization and integration of antecedent DFAS governance concepts into an AML-specific decision-governance doctrine were independently undertaken by the author, while prior DFAS constructs were treated as intellectual lineage rather than external evidence of the identified governance gap.
Narrative Structure and Analytical Exposition	Class 2	AI assistance was limited to controlled linguistic refinement, clarity, and structural presentation.
Tables, Figures, and Conceptual Presentation	Class 2	The substantive governance mechanisms and relationships were author-designed; AI assistance was restricted to formatting and presentation refinement.
Literature Review and Cross-Referencing	Class 2	Literature interpretation, evidence selection, and theoretical integration were author-directed; AI assistance was limited to citation and reference alignment.
Appendices and Compliance Documentation	Class 2	AI assistance was restricted to formatting and structural consistency.

Overall Manuscript Classification: Class 2 – AI-Assisted, Author-Validated

DFAS-AML Doctrine, Five Governance Mechanisms, Evidence-to-Design Architecture, Decision-Governance Sequence, and Conceptual Evaluation Architecture: Class 1 – Fully Human-Authored

Additional Clarifications

Human-Originated Content (Class 1): All substantive contributions including DFAS-AML, the Evidence-to-Design Governance Architecture, the five-component governance architecture, A-AMLEEF, A-AMLOCC, A-AMLD, A-SCEAML, A-AMIL, the five derived governance requirements, the integrated AML decision-governance sequence, the recursive governance logic, and the six-criterion conceptual evaluation architecture were independently conceptualized, structured, and developed by the author. The manuscript explicitly positions DFAS-AML as an AML-specific governance architecture integrating functions across the complete decision lifecycle rather than treating governance as limited to the detection model. DFAS-AML

AI-Assisted Content (Class 2): AI tools were used only for non-substantive linguistic refinement, formatting, presentation, and citation alignment.

Integrity Safeguards: All AI-assisted content was manually reviewed and validated under DFAS-FEP v1.6, with authorship classification and traceability documentation maintained.

Appendix B – Version History Log (VHL)

Version 1.0 – First submitted to SSRN in July 27, 2025, Published in July 30, 2025.

SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5367894

- The initial manuscript established the core conceptual architecture of DFAS-AML, including A-AMLD, A-AMLEEF, A-AMLOCC, A-SCEAML, and A-AMIL, their governance functions, decision logic, integrated workflow, and illustrative application scenarios. These substantive conceptual elements were independently developed by the author.
- AI tools were used only for non-substantive assistance, including linguistic refinement, structural organization, formatting, and presentation. The conceptual contribution, analytical reasoning, interpretation, and final validation remained under the author's control.
- Alaali-ACS: 0 | Classification: Class II – AI-Assisted, Author-Validated Version 2.0 – Substantially Revised in Response to the Journal's Major Revision Decision, September 26, 2026.
- The manuscript's substantive conceptual architecture—including DFAS-AML, the Evidence-to-Design Governance Architecture, the five governance requirements of Explanation Governance, Accountable Override, Ethical Resilience, Sovereign Calibration, and Traceability, and their integration through A-AMLEEF, A-AMLOCC, A-AMLD, A-SCEAML, and A-AMIL—was fully human-authored. The integrated decision-governance sequence, recursive governance logic, AML-specific specialization of antecedent DFAS governance concepts, and conceptual evaluation architecture were independently conceptualized and developed by the author. The manuscript distinguishes conceptual construction and evaluation from empirical validation and does not claim established operational effectiveness. DFAS-AML
- AI tools were used exclusively for non-substantive support functions, limited to controlled

linguistic refinement, structural presentation, formatting, and citation/reference alignment, in compliance with DFAS-FEP v1.6 presentation standards.

- Alaali-ACS: 0 | Classification: Class II – AI-Assisted, Author-Validated