

From Pilots to Production: Enterprise Genai Integration Across Media and Financial Services

Ashish Kalaskar

Department Of Computer Science, Artificial Intelligence & Machine Learning, MIT AOE, India

doi.org/10.64643/ijirt.209085-459

Abstract—Despite corporate investments exceeding \$30–40 billion in Generative Artificial Intelligence (GenAI), enterprise deployments exhibit a profound failure rate. As documented in the landmark MIT NANDA *State of AI in Business 2025* report, 95% of enterprise AI pilots fail to yield measurable P&L return. The root cause is rarely model capability; rather, it stems from a systemic "learning and integration gap"—a failure to architect AI agents directly into core Enterprise Systems-of-Record (SoR) with automated guardrails, context-grounded telemetry, and closed feedback memory loops

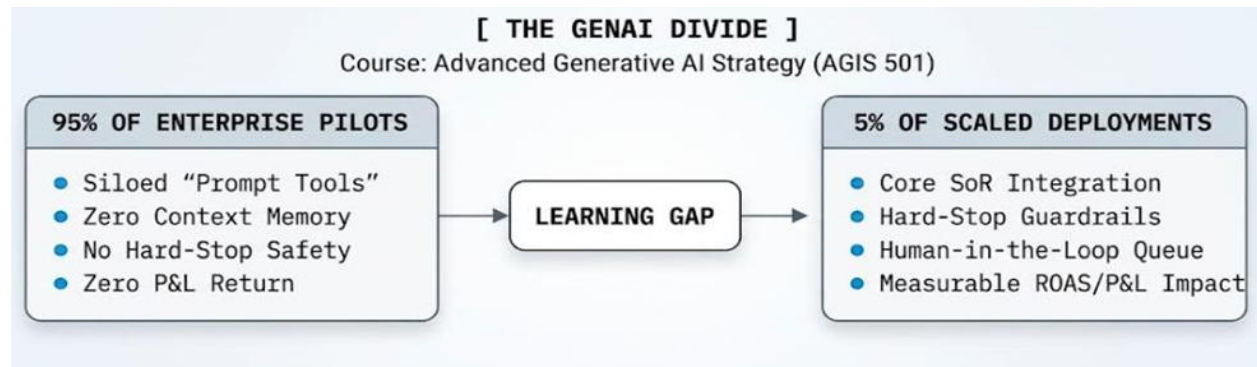
This research paper presents a blueprinting framework designed to cross the "GenAI Divide" by analyzing two starkly contrasting enterprise paradigms: Media, Marketing & Advertising (characterized by high velocity, rapid iteration, and real-time bid optimization) and Retail & Commercial Banking (characterized by rigid regulatory governance, zero-trust data privacy,

and heavy financial liability). By conducting a workflow-by-workflow architectural audit across eight high-impact sub-functions, this paper establishes the structural protocols, governance mechanisms, and human-in-the-loop (HITL) escalations required to transition brittle "Shadow AI" prompts into scalable, enterprise-grade AI systems.

I. INTRODUCTION

1.1. The Enterprise GenAI Divide

The enterprise AI landscape in 2026 is defined by a paradoxical disconnect. While individual employee adoption of consumer-facing models ("Shadow AI") exceeds 90%, formal enterprise GenAI initiatives remain heavily stalled



AI scaling succeeds not through raw model capacity, but by embedding agentic models directly into operational pipelines via context retrieval, deterministic guardrails, and continuous learning loops rather than treating them as isolated generative interfaces

1.2. Industry Selection Rationale: High Velocity vs. High Governance

To establish a universally applicable framework, this research focuses on two sectors that represent opposite

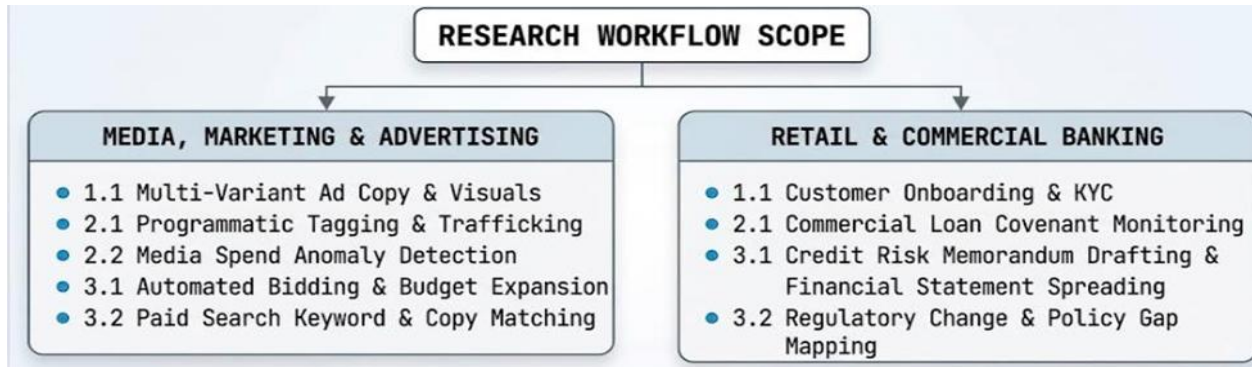
ends of the operational risk and velocity spectrum:

- **Media & Marketing (High Velocity):** Operates on rapid execution (sub-second to hours), un-structured asset generation, and real-time ad bidding, where key risks include ad-spend waste, brand safety, and algorithmic bias.
- **Financial Services (High Governance):** Operates under zero-trust, strict regulatory compliance (onboarding, credit memoranda, covenant tracking), where critical risks center on regulatory fines, fiduciary liability, and PII exposure.

1.3. Scope of Investigation

This paper systematically examines eight core sub-

workflows across these two sectors:



II. RESEARCH METHODOLOGY & ANALYTICAL FRAMEWORK

2.1. Dual-Taxonomy Blueprinting Methodology

For each selected workflow, the research applies a standardized diagnostic structure that decomposes the operational problem from immediate failure points to system-of-record integration:

- Sub-Function & Operational Objective: Pinpointing the precise enterprise task and business owner.
- Problem & Bottleneck Identification: Defining why human execution fails at scale (e.g., latency, cognitive overload, manual error rate) and why basic AI prompts fail (e.g., hallucinations, context loss).
- Task Complexity & Frequency: Quantifying runtime requirements (real-time continuous vs. batch processing) and algorithmic complexity.
- Data Requirements & Sensitivity Matrix: Classifying necessary input streams (CRM, pixels, financial ledgers, legal documents) and privacy risks (PII, sensitive corporate financial data, GDPR/CCPA, SEC/FINRA compliance).
- Systems-of-Record (SoR) & Integration Architecture: Mapping mandatory integrations across enterprise software (e.g., Salesforce, Segment, Google Ads API, nCino, Moody's, Core Banking Systems).

III. MEDIA, MARKETING, AND ADVERTISING INDUSTRY WORKFLOW MAPPING & EVIDENCE BASE

This document establishes the empirical evidence base and workflow analysis for the Media, Marketing, and Advertising industry.

3.1. Creative and Content Production

Workflow 1.1: Multi-Variant Ad Copy & Visual Asset Generation

- Sub-function: Creative and content production
- Workflow: Rapid generation and resizing of digital ad copy, banner graphics, and social display variations across multi-channel campaigns.
- Current problem: Creative teams spend over 60% of their bandwidth on manual asset formatting, copy variations, and aspect ratio adjustments rather than high-level campaign concepts. This bottleneck delays campaign launches and limits multivariate A/B testing capacity.
- Task frequency and complexity: High frequency (daily/weekly per campaign sprint); Low-to-Medium cognitive complexity (formatting, re-skinning, minor copy tweaks).
- Data required and sensitivity: Brand guidelines, historical ad performance data, product catalogs, stock imagery, copy guidelines. *Sensitivity*: Low-to-Medium (pre-launch campaign copy and unreleased product assets are commercially sensitive prior to launch).
- Existing systems and integration needs: Adobe Creative Cloud (Photoshop, Illustrator), Figma, Digital Asset Management (DAM) systems (e.g., Bynder, Brandfolder), CMS, and ad server asset repositories.
- Risks:
 - *Copyright*: High risk of generative image models training on copyrighted visual styles or incorporating non-licensed source elements.
 - *Privacy*: Low risk, provided personal consumer data is not injected into creative prompts.

- *Bias*: High risk of perpetuating visual stereotypes in generated human model assets.
- *Misleading claims*: medium risk of AI hallucinatory copy making unsubstantiated product claims.
- *Brand safety*: High risk of off-brand tone, distorted logos, or artifact-heavy visuals.

AI tool type: Tool Stack Architecture

- Enterprise Stack: Enterprise Brand RAG + Parametric Engine
- *Generative Layer*: Adobe Firefly Services API or Jasper Enterprise (enforcing brand voice, style vectors, and copyright indemnification).
- *System of Record*: Connected via API to Digital Asset Management (DAM) (e.g., Bynder, Brandfolder) and Figma.
- *Orchestration Layer*: Low-code workflow tools (e.g., Vellum or n8n) to route prompts, parameters, and outputs.

Standard Operating Procedure (SOP)

1. Trigger & Context Retrieval: Copywriter or Designer submits a campaign brief in Figma/DAM. The system automatically pulls approved brand guidelines, visual style kits, and historical high-performing creative vectors from the DAM.
2. Parametric Draft Generation: The API generates multi-aspect ratio visual assets and copy variations constrained strict-bound style tokens (fonts, hex codes, non-licensed visual exclusions).
3. HITL Review Gate: Senior Art Director receives the generated batch in a review queue.
 - *Pass*: Director approves.
 - *Reject/Edit*: Director modifies copy/layout. Crucial NANDA Loop: Edits are logged to a vector feedback store to refine future generation parameters.
4. Deployment & Telemetry: Approved assets auto-tag and push to the ad server. Post-launch CTR and conversion telemetry flow back into the DAM to score creative vector effectiveness.
 - Required human review and escalation: Mandatory Senior Art Director and Copywriter approval for all AI-generated masters before asset distribution; escalation to Brand Legal if synthetic human models are generated for commercial ads.

- Possible success metrics:
 - *Reduction* in creative production lead time (target: >50% reduction).
 - *Cost* per asset variation produced.
 - *Click-through rate* (CTR) uplift across multivariate creative sets.
- Research evidence needed before making a recommendation:
 - *Sourced Fact*: Gartner's CMO Spend Survey reveals that 77% of GenAI-using marketers apply it to creative development, making it the most widespread use case. Gartner also notes GenAI reduces content production costs by up to 60%.
 - *Provisional Assumption*: Adopting brand-governed visual generation tools will reduce agency visual adaptation outsourcing costs by 35–45% without degrading brand equity metrics.

3.2. Advertising Operations (Ad Ops)

Workflow 2.1: Programmatic Campaign Tagging, Ad Trafficking, and Taxonomies Verification

- Sub-function: Advertising operations
- Workflow: Generating tracking URLs (UTMs), placing ad tags/pixels, configuring ad server flight dates, and verifying naming conventions across platforms.
- Current problem: Ad trafficking involves tedious manual data entry across multiple DSPs, ad servers, and verification tools. Human error in UTM parameters breaks attribution models and wastes ad dollars.
- Task frequency and complexity: Daily per campaign deployment; High detail-oriented operational complexity.
- Data required and sensitivity: Campaign brief taxonomies, landing page URLs, ad server specs, client tracking macros. *Sensitivity*: Medium (pre-launch campaign structure and pricing details).
- Existing systems and integration needs: Ad servers (Google Ad Manager, Flashtalking), DSPs (The Trade Desk, DV360), Verification tools (DoubleVerify, IAS), Analytics (Google Analytics 4, Adobe Analytics).
- Risks:
 - *Copyright*: N/A.
 - *Privacy*: Medium risk (improper placement of retargeting pixels on sensitive data pages violating consent rules).

- *Bias: Low risk.*
- *Misleading claims: Low risk.*
- *Brand safety: Low risk.*
- *Compliance: High risk (non-compliance with consent management frameworks like TCF v2.2).*
- *Financial risk: High (broken tracking leads to unmeasurable ad spend and inaccurate billing).*

AI tool type

- Enterprise Stack: Document AI + Anomaly Detection Orchestration
- *Automation Core:* Microsoft Power Automate or n8n developer automation platform.
- *Validation Engine:* Custom Python / Unsupervised ML anomaly detection models running on AWS Bedrock AgentCore or Google Cloud Vertex AI.
- *System of Record:* Direct integration with Mediaocean, Google Ad Manager, and DSP APIs.

Standard Operating Procedure (SOP)

1. Automated Ingestion: The agent automatically extracts campaign briefs, ad server flight logs, UTM taxonomies, and publisher invoices via API.
2. Automated Taxonomy & Tagging Verification: The Document AI runs a 100% audit on tracking tags, macro formatting, and flight dates prior to launch.
3. Anomaly & IVT Detection: Machine learning algorithms continuously monitor impression delivery pacing and Invalid Traffic (IVT) thresholds across DSPs.

4. Escalation & Dispute Procedure:

- *Normal Pacing: System auto-reconciles end-of-month delivery against DSP invoice data.*
- *Pacing Anomaly (>10% variance or >3% IVT):* System immediately pauses campaign flighting and generates a dispute ticket sent directly to the Ad Ops Lead and Publisher Rep.
- Required human review and escalation: Automated pre-flight validation checklist with human sign-off on 10% sample audit; mandatory escalation if pixel verification fails automated ping tests.
- Possible success metrics:
 - *Zero tracking parameter error rate.*
 - *Trafficking time per campaign reduced by >80%.*
 - *On-time campaign launch compliance rate (100%).*

- Research evidence needed before making a recommendation:
- *Sourced Fact:* IAB State of Data 2025 reports that only 30% of organizations have fully integrated AI across the media campaign lifecycle, citing ad ops fragmentation as a primary barrier.
- *Provisional Assumption:* Automated tag verification scripts will eliminate attribution loss caused by broken UTM links and misconfigured pixel triggers.

Workflow 2.2: Media Spend Reconciliation and Delivery Anomaly Detection

- Sub-function: Advertising operations
- Workflow: Continuous monitoring of ad impression pacing, ad fraud rates, viewability scores, and cross-platform invoice reconciliation.
- Current problem: Media planners spend days at end-of-month reconciling platform billing discrepancies, identifying over-delivery/under-delivery, and disputing invalid traffic (IVT) charges manually.
- Task frequency and complexity: Continuous monitoring / Monthly reconciliation; High numerical complexity.
- Data required and sensitivity: DSP spend reports, ad server delivery logs, publisher invoices, verification reports (IAS/DV). *Sensitivity:* High (financial expenditure data and contractual CPM rates).
- Existing systems and integration needs: Media ERPs (Mediaocean, Prisma), financial accounting software (SAP, NetSuite), DSP/Ad Server APIs.
- Risks:
 - *Copyright: N/A.*
 - *Privacy: Low risk.*
 - *Bias: Low risk.*
 - *Misleading claims: Low risk.*
 - *Brand safety: High risk (undetected ad delivery on fraud/MFA—Made for Advertising—websites).*
 - *Compliance: High risk (financial reporting accuracy and auditability).*
 - *Financial risk: High (paying for non-human traffic or failing to claw back under-delivered media budgets).*

Possible AI tool type Recommended tools

1. Double Verify Media Advantage Platform current cloud SaaS release Purpose: Independently measures invalid traffic, viewability, brand

suitability, and delivery quality across digital advertising placements. DoubleVerify Media AdVantage

2. Mediaocean Prisma Finance current cloud SaaS release

Purpose: Matches media orders to vendor invoices, automates billing reconciliation, and integrates media-finance records with ERP and accounting systems. Prisma Finance

Procedure: Media Spend Reconciliation and Delivery Anomaly Detection

1. The Ad Ops team imports approved campaign budgets, insertion orders, CPMs, pacing targets, delivery commitments, and make-good terms into Prisma Finance.
2. DSPs, ad servers, publishers, and verification platforms send daily delivery, spend, impression, viewability, and invalid-traffic data through approved APIs.
3. DoubleVerify independently measures invalid traffic, viewability, geography, and brand-suitability signals for eligible campaign inventory.
4. The Ad Ops dashboard compares planned delivery, ad-server delivery, DSP spend, DoubleVerify media-quality results, and publisher invoices at campaign, publisher, placement, and creative level.
5. The system flags pacing deviations, unexpected spend increases, over-delivery, under-delivery, invalid-traffic spikes, low viewability, and mismatches between delivered impressions and invoiced impressions.
6. The media planner investigates alerts using the underlying platform logs and verification evidence. The AI system may prioritize anomalies but does not automatically pause campaigns, dispute invoices, or approve payments.
7. Confirmed invalid-traffic, delivery, or billing discrepancies are recorded in Prisma Finance and sent to the publisher, DSP, or ad network as a documented credit, make-good, or invoice-dispute request.
8. Finance reviews and approves material invoice variances according to the organization's approved financial-control thresholds before payment is released.
9. The pilot is evaluated using reconciliation-close time, invoice-match rate, verified invalid-traffic recovery, over-spend incidents, unresolved

discrepancy value, and number of manual reconciliation hours.

- Required human review and escalation: Immediate human alert trigger if daily pacing strays >10% off trajectory or IVT spikes above 3%; Finance team review required for invoice variance >\$500.
- Possible success metrics:
 - Days to close monthly media reconciliation (target: reduced from 10 days to 1 day).
 - Recovery percentage of invalid traffic (IVT) funds from publishers.
 - Elimination of campaign over-spend penalties.
- Research evidence needed before making a recommendation:
 - *Sourced Fact*: According to Vention research, operational process transformation using AI yields a 25–31% labor efficiency improvement and 5–10% cost reduction.
 - *Provisional Assumption*: Real-time ML pacing anomaly detection will reduce wasted ad spend on underperforming or fraudulent placements by 8–12%.

3.3. Performance Marketing

Workflow 3.1: Automated Campaign Bidding, Budget Allocation & Audience Expansion

- Sub-function: Performance marketing
- Workflow: Real-time optimization of advertising auction bids, budget shifting across channels (Search, Social, Display, Retail Media), and lookalike audience discovery.
- Current problem: Human media buyers cannot process millions of auction signals per second across fragmented platforms (Google, Meta, Amazon, TikTok), leading to sub-optimal budget allocation, overbidding, and budget waste.
- Task frequency and complexity: Real-time / Continuous; High algorithmic complexity.
- Data required and sensitivity: Conversion tracking data, pixel telemetry, first-party CRM data (customer LTV, email hashes), ad platform APIs, cost per acquisition (CPA) targets. *Sensitivity*: High (contains PII and proprietary revenue figures).
- Existing systems and integration needs: Ad platforms (Google Ads, Meta Ads Manager, Amazon Ads), Customer Data Platforms (CDPs like

Segment, Tealium), Multi-Touch Attribution (MTA) tools.

- Risks:
 - *Copyright*: N/A.
 - *Privacy*: High risk (improper processing of customer PII across ad networks violating GDPR/CCPA).
 - *Bias*: High risk (algorithmic audience exclusion leading to discriminatory ad targeting in housing, employment, or credit).
 - *Misleading claims*: N/A.
 - *Brand safety*: medium risk (bidding algorithms placing ads on low-quality or fraudulent publisher inventory).
 - *Compliance*: High risk (privacy law breaches, failure to honor opt-out signals).
 - *Financial risk*: Very High (runaway automated bidding algorithms burning budget without delivering conversions).

Possible AI tool type: 1. Google Ads Smart Bidding current platform-native AI Purpose:

Automates auction-level bid optimization using machine learning, conversion/value signals, and contextual information. Google states that Smart Bidding sets bids for individual auctions and supports objectives such as Target CPA and Target ROAS.

2. Skai current cross-channel AI optimization platform Purpose:

Use as the cross-channel decision and optimization layer across paid-media environments rather than attempting to replace every platform's native bidding model. It can be evaluated for budget allocation, campaign optimization, and cross-channel performance comparison

Procedure: Automated Campaign Bidding, Budget Allocation & Audience Expansion

1. Define business objectives and guardrails.
Set campaign-level objectives such as CPA, ROAS, conversion volume, and customer-value targets. Establish hard limits for daily spend and the proposed 15% maximum automated daily budget shift.

2. Connect approved first-party and advertising data.
Integrate conversion data, approved CRM/LTV signals, campaign performance data, and advertising-platform APIs. Customer data should only be

transferred where the organization's consent, privacy, contractual, and platform-policy requirements are satisfied. Google, for example, has specific policies governing Customer Match and first-party data use.

3. Use native platform AI for auction-level bidding.
Configure Google Ads Smart Bidding around the organization's selected CPA or ROAS objective. Its advantage is that it operates directly within the auction and can evaluate contextual signals at auction time.

4. Use the cross-channel layer for portfolio-level decisions.

Feed campaign-level performance from Google, Meta, Amazon, TikTok, and other approved channels into the cross-channel optimization system. Compare marginal performance, pacing, CPA/ROAS, and conversion value across channels before recommending budget movements.

5. Start with recommendations rather than unrestricted autonomous allocation.

During the pilot, the system should recommend budget reallocations and audience opportunities. A human performance-marketing lead approves material changes. This follows the research principle that organizations should begin with a narrow, valuable workflow and learn through user feedback rather than immediately automating the entire process.

6. Introduce controlled automation after validation.

Once performance and data quality are validated, permit automated budget movements within predefined limits—for example, no more than 15% of the daily budget without human approval. Larger reallocations require sign-off.

7. Monitor for model and business anomalies.

Trigger immediate human review when CPA/ROAS deviates materially from the approved range, spend accelerates unexpectedly, conversion tracking changes, or audience performance shows unusual demographic disparities.

8. Evaluate native-vs-cross-channel performance experimentally.

Do not assume that a cross-channel model will outperform Google/Meta/Amazon's native algorithms. Run a controlled test comparing the existing platform-

native strategy against the proposed cross-channel allocation approach.

9. Keep sensitive audience expansion constrained.

Audience expansion should exclude prohibited or legally sensitive targeting categories and respect consent/opt-out requirements. Human review remains mandatory where automated targeting could create discriminatory outcomes.

10. Scale only after measurable evidence.

Compare ROAS, CPA, budget utilization, incremental conversions, wasted spend, and manual optimization hours against a pre-defined baseline. Scale only if the system demonstrates sustained improvement without increasing privacy, compliance, or financial risk

- Required human review and escalation: Hard stop caps on automated budget reallocation (e.g., maximum 15% daily shift without human sign-off); daily audit of CPA/ROAS anomalies by Performance Marketing Directors.
- Possible success metrics:
 - *Return* on Ad Spend (ROAS) improvement.
 - *Reduction* in Cost Per Acquisition (CPA).
 - *Budget* utilization efficiency (zero unspent or overspent pacing errors).
- Research evidence needed before making a recommendation:
 - *Sourced Fact*: Automated bidding and AI-driven campaign optimizations dominate paid media channels, with over 80% of digital ad spend managed via platform-native machine learning models.
 - *Provisional Assumption*: Cross-channel predictive AI allocation models will outperform platform-native algorithms (e.g., Meta Advantage+ / Google Performance Max) by preventing single-ecosystem bias.

Workflow 3.2: Paid Search Keyword Expansion & Search Intent Copy Matching

- Sub-function: Performance marketing
- Workflow: Discovering long-tail search keywords, analyzing searcher intent, and dynamically generating semantically aligned landing page/search

ad copy combinations.

- Current problem: Manual keyword research misses fast-emerging consumer query trends. Static search ads often fail to tightly match evolving search intent, lowering Google Quality Scores and driving up Cost-Per-Click (CPC).
- Task frequency and complexity: Continuous/Weekly; Medium complexity.
- Data required and sensitivity: Search volume metrics, Google Search Console data, SEMrush/Ahrefs data, historical keyword conversion rates. *Sensitivity*: Low-to-Medium (competitive strategy data).
- Existing systems and integration needs: Google Ads API, Microsoft Advertising API, SEO/SEM suites, CMS for dynamic landing page generation.
- Risks:
 - *Copyright*: Low risk.
 - *Privacy*: Low risk.
 - *Bias*: Low risk.
 - *Misleading claims*: High risk (AI copy making incorrect product/pricing assertions in dynamic ad text).
 - *Brand safety*: medium risk (bidding on negative or brand-damaging search queries).
 - *Compliance*: Medium risk (violating trademark policies of competitors or regulated product disclaimers).
 - *Financial risk*: Medium (buying irrelevant high-cost search traffic due to hallucinated intent classification).

Possible AI tool type: 1. Google Ads Keyword Planner current Google Ads platform tool:

Use it as the primary keyword-discovery and validation layer. It can generate new keyword ideas, provide monthly search estimates, historical trends, and bid estimates, allowing AI-generated keyword candidates to be validated against actual search-market data rather than relying only on an LLM's inferred intent.

2. Microsoft Advertising AI Recommendations & Keyword Planner current platform-native AI:

Use it as a secondary search-intelligence and optimization layer for keyword expansion, trending queries, ad-group suggestions, and keyword recommendations. Microsoft states that its

recommendations use historical account performance, campaign settings, and trends, while Keyword Planner provides search-volume trends and cost estimates.

Procedure: Paid Search Keyword Expansion & Search Intent Copy Matching

1. Collect the existing search-term and keyword baseline.

Import historical Google Ads/Microsoft Ads keywords, search-query reports, conversion rates, CPC, CTR, landing pages, and existing negative-keyword lists.

2. Generate new keyword candidates.

Use Google Keyword Planner to discover related and long-tail keywords. Validate candidates using search-volume, trend, and bid information rather than allowing an LLM to invent keywords independently.

3. Use Microsoft Keyword Planner as a second evidence source.

Compare keyword suggestions, historical search trends, competition, and estimated costs. This provides an additional signal before adding emerging keywords to campaigns.

4. Classify search intent.

Apply an NLP/LLM layer to classify candidate queries into categories such as informational, navigational, commercial investigation, and transactional. The model should provide a confidence score rather than treating its classification as certain.

5. Filter candidates using business evidence.

Prioritize keywords that combine relevant search intent with sufficient search demand, acceptable CPC, historical conversion performance, and a clear connection to an existing product or landing page.

6. Build intent-matched ad groups.

Group semantically related queries so that each ad group has a coherent search intent and landing-page destination. Microsoft similarly recommends organizing related keywords into highly related ad groups and using negative keywords to prevent irrelevant traffic.

7. Generate ad-copy variants with controlled inputs.

The LLM may generate headlines/descriptions based

only on an approved product-information database, pricing data, offers, disclaimers, and landing-page content. It should not invent prices, features, guarantees, or claims.

8. Match copy to landing-page intent.

Each generated ad should be checked against the destination page to ensure that the promise made in the advertisement is actually supported by the landing page.

9. Apply negative-keyword and brand-safety controls. SEM specialists review the proposed negative-keyword list weekly. Queries involving irrelevant, damaging, prohibited, or misleading contexts should be excluded before activation. Negative keywords are specifically designed to prevent ads from appearing for irrelevant searches.

10. Run a controlled experiment rather than assuming AI improves conversion.

Compare AI-assisted keyword expansion and intent-matched copy against the existing/static approach. Measure incremental conversions, CTR, CPC, conversion rate, Quality/Ad relevance indicators, and wasted spend.

11. Keep human approval for new campaigns and claims.

AI can recommend keywords and generate copy, but the SEM specialist approves new keyword clusters, negative keywords, landing-page matches, and claims before publication.

12. Continuously feed performance back into the system.

Weekly performance data should update the keyword and intent model. Poor-performing or irrelevant terms are removed, while validated emerging queries can be promoted into the active keyword strategy.

- Required human review and escalation: Negative keyword exclusion lists managed weekly by SEM specialists; automated rule triggers that pause ads with CTRs dropping below baseline.
- Possible success metrics:
 - *Incremental* conversion volume from long-tail keywords.
 - Google Ad Quality Score improvements (target:

>8/10 average).

- *CPC* reduction resulting from improved ad relevance.
- Research evidence needed before making a recommendation:
 - *Sourced Fact*: Modern performance marketing relies heavily on predictive audience and search intent algorithms, with AI-driven optimizations driving 10–20% conversion rate increases.
 - *Provisional Assumption*: LLM-generated search copy tied dynamically to search intent will increase conversion rates by 15% compared to static template-based ad groups.

IV. FINANCIAL SERVICES AND INSURANCE (FSI) WORKFLOW MAPPING & EVIDENCE BASE

Applying the empirical lessons of the MIT NANDA (2025) report, the Financial Services and Insurance (FSI) sector represents both the highest potential value pool for Generative AI and the highest operational risk environment.

While McKinsey estimates GenAI could create \$200 billion to \$340 billion in annual value across global banking (equivalent to 9–15% of operating profits), FSI organizations face strict regulatory mandates (SEC, FINRA, OCC, CFPB, NAIC, GDPR) that punish black-box hallucinations, data leaks, and algorithmic bias.

4.1. Retail Banking

Workflow: Customer Onboarding & KYC Identity Verification Assistance

- Sub-function: Retail Banking & Wealth Management
- Workflow: Extracting, validating, and cross-referencing customer identity documents (passports, driver's licenses, utility bills) during digital account opening.
- Current problem: Manual document verification creates severe friction during onboarding, leading to high drop-off rates (often 40–50%) and lengthy approval cycles (3–5 business days).
- Task frequency and complexity: Real-time / High-volume daily; Medium technical and regulatory complexity.

- Data required and sensitivity: Extremely High Sensitivity. Biometric data, national identity numbers (SSN/Tax ID), proof of address, facial photographs.
- Existing systems and integration needs: Customer Onboarding Platforms (Appway, Pega), Identity Verification (Jumio, Onfido), Core Banking APIs, Sanctions/PEP Screening databases.
- Risks:
 - *Copyright*: Low risk.
 - *Privacy*: Extremely High risk (mishandling biometric data or storing unencrypted PII).
 - *Bias*: High risk (computer vision models exhibiting higher false rejection rates on non-majority demographic groups).
 - *Misleading claims*: Low risk.
 - *Brand safety*: medium risk (frustrating legitimate customers through false fraud flags).
 - *Compliance*: Very High risk (failing FinCEN Customer Due Diligence rules or GDPR/CCPA consent mandates).
 - *Financial risk*: High (fines for onboarding sanctioned individuals or fraudulent identity theft accounts).

AI tool type.: Recommended tools

1. Entrust Identity Verification SDK 2026 major release

Purpose: Captures and verifies identity documents, performs biometric/liveness checks, and routes KYC outcomes through configurable onboarding workflows. Entrust SDK documentation

2. Jumio Identity Verification cloud platform (current SaaS release)

Purpose: Verifies government IDs, selfies/liveness, and supporting documents such as proof of address, then produces a fraud-risk decision for the bank's review process. Jumio Identity Verification procedure

1. Define the risk policy first

Compliance, fraud, privacy, and operations teams must decide which cases can be automated, which require review, and which require rejection.

2. Use risk-based routing

Low-risk, high-confidence cases may proceed automatically. Any mismatch, suspected fraud, failed

liveness check, sanctions/PEP match, unsupported document, or repeated attempt goes to human review.

3. Do not use “more than 5% discrepancy”

Replace it with vendor evidence signals and bank-approved risk thresholds. A document check is normally pass/fail or risk-scored; a fixed 5% rule is unclear and difficult to justify to auditors.

4. Keep biometric data tightly controlled

Define consent, encryption, access roles, data residency, retention, deletion, audit logs, and vendor-subprocessor rules before the pilot.

5. Test fairness and false-rejection rates

Measure outcomes across document types, countries, age groups, lighting conditions, and—where lawful and appropriate—demographic groups. A low fraud rate is not sufficient if legitimate customers are rejected disproportionately.

6. Pilot in one customer segment

Start with low-risk retail accounts, one country or jurisdiction, and a limited set of identity documents. Do not launch globally first.

7. Require human compliance review for high-risk cases

The AI tool may flag and prioritize; it must not make final AML, sanctions, or suspicious-activity decisions independently.

8. Scale only after evidence

Expand when the pilot improves onboarding time and completion without increasing fraud losses, compliance failures, complaints, or false rejections

- Required human review and escalation: Automatic escalation to Compliance/AML Operations for any document flagged with >5% discrepancy or high-risk PEP/sanctions match.
- Possible success metrics:
 - *Straight-Through Processing* (STP) onboarding rate (target: >80%).
 - *Abandonment* rate during digital account opening.
 - *Time-to-active* account creation (reduced from days to minutes).

- Research evidence needed before making a recommendation:

- *Sourced Fact*: IDP market benchmarks indicate automated document ingestion reduces manual onboarding processing time by up to 70% while improving data extraction accuracy.
- *Provisional Assumption*: Replacing legacy optical character recognition (OCR) with multimodal LLM vision tools will reduce human-in-the-loop manual review volume by 50% for unreadable documents.

4.2. Commercial & Corporate Lending

Workflow 2.1: Credit Risk Memorandum Drafting & Financial Statement Spreading

- Sub-function: Commercial & Corporate Lending
- Workflow: Extracting financial statements, tax returns, and balance sheets from commercial loan applicants, spreading financial ratios, and drafting comprehensive credit risk memos.
- Current problem: Commercial credit analysts spend 15–20 hours per loan applicant manually re-keying audited financial statements and drafting repetitive credit narrative memos, delaying loan approval turnaround.
- Task frequency and complexity: Weekly per deal pipeline; High analytical and financial complexity.
- Data required and sensitivity: High Sensitivity. Corporate financial statements, tax returns, debt schedules, bank statements, industry benchmarks.
- Existing systems and integration needs: Commercial Loan Origination Systems (nCino, Moody’s CreditLens), Accounting Software APIs, Financial Spreading Software, Enterprise Data Warehouses.
- Risks:
 - *Copyright*: Low risk.
 - *Privacy*: Medium-to-High risk (confidential corporate tax filings and unannounced financial results).
 - *Bias*: Medium risk (historical lending bias embedded in automated credit risk scoring).
 - *Misleading claims*: High risk (AI hallucinating financial ratios, miscalculating debt service coverage, or understating credit risk).
 - *Brand safety*: medium risk.
 - *Compliance*: Very High risk (Equal Credit Opportunity Act / ECOA, Fair Lending disclosures, OCC credit risk guidelines).

- *Financial risk*: Extremely High (approving bad loans due to faulty AI financial ratio analysis leading to loan default).

Possible AI tool type: nCino Automated Spreading current cloud SaaS release Purpose: Extracts, maps, reconciles, and analyzes financial data from tax returns, audited statements, 10-Ks, 10-Qs, and company-prepared financial statements within the commercial-lending workflow. nCino Automated Spreading

- Moody's Lending Suite Loan Origination current cloud SaaS release

Purpose: Automates financial spreading and credit scoring and produces a GenAI-assisted, editable first-draft credit memorandum from verified lending data. Moody's Loan Origination

Procedure: Credit Risk Memorandum Drafting and Financial Statement Spreading

1. The relationship manager submits the borrower's financial statements, tax returns, debt schedules, bank statements, and loan request through the bank's approved loan-origination system.
2. The selected platform ingests documents and extracts financial data into bank-approved spreading templates.
3. The system identifies missing fields, conflicting values, low-confidence extractions, unsupported document formats, and unusual changes from prior reporting periods.
4. The platform calculates debt-service coverage, leverage, liquidity, profitability, covenant headroom, and other ratios only through bank-approved calculation formulas.
5. The system generates a first-draft credit memorandum using verified spread data, approved credit-policy content, borrower history, and documented risk factors.
6. The credit analyst reconciles every extracted financial value, ratio, and narrative statement against the original borrower documents.
7. The analyst records corrections, exceptions, and unresolved data issues in the loan file. These corrections must not be silently incorporated without an auditable record.
8. The senior credit officer reviews and signs the final credit memorandum, loan structure, risk rating, and

approval recommendation. The AI system does not approve, reject, price, or risk-rate the loan independently.

9. The pilot is evaluated using memo-preparation time, extraction accuracy, calculation-error rate, manual-review volume, approval-to-funding cycle time, cost per funded loan, and fair-lending review outcomes.

- Required human review and escalation: Mandatory sign-off by Senior Credit Officer; AI-generated credit memos are treated strictly as "first drafts" requiring 100% human audit of numbers against original tax filings.

- Possible success metrics:

- *Credit* memo drafting turnaround time (McKinsey benchmark target: 20–60% productivity gain).
- *Commercial* loan approval-to-funding cycle time.
- *Cost* per funded loan.

- Research evidence needed before making a recommendation:

- *Sourced Fact*: McKinsey reports that AI-assisted credit memo generation in commercial banking delivers 20% to 60% productivity gains and cuts decision turnaround times by ~30%.
- *Provisional Assumption*: Automated financial statement spreading will achieve >99% mathematical precision when paired with structured table-parsing models, eliminating manual spreadsheet re-keying.

Workflow 2.2: Commercial Loan Documentation & Covenant Compliance Monitoring

- Sub-function: Commercial & Corporate Lending
- Workflow: Tracking corporate borrower compliance against debt covenants (e.g., leverage ratios, liquidity thresholds, reporting deadlines) extracted from complex loan agreements.
- Current problem: Commercial loan portfolios contain thousands of bespoke credit agreements. Portfolio managers manually review compliance certificates and quarterly reports, often missing early covenant breaches.
- Task frequency and complexity: Quarterly / Continuous; High legal and financial complexity.
- Data required and sensitivity: High Sensitivity. Executed credit agreements, compliance certificates, quarterly SEC filings / financial statements.

- Existing systems and integration needs: Loan Administration Platforms (ACBS, LoanIQ), Contract Lifecycle Management (CLM) systems, Risk Management Dashboards.
- Risks:
 - *Copyright*: Low risk.
 - *Privacy*: High risk (unreleased commercial borrower earnings data).
 - *Bias*: Low risk.
 - *Misleading claims*: High risk (misinterpreting complex legal loan definitions, e.g., EBITDA adjustments).
 - *Brand safety*: medium risk.
 - *Compliance*: High risk (OCC loan portfolio rating standards).
 - *Financial risk*: High (unidentified borrower distress resulting in uncollateralized defaults).

Possible AI tool type: Domain-specific Legal NLP / Document Understanding Model paired with financial calculation rule Recommended tools

1. Moody's Lending Suite Monitoring Automated Covenants, current cloud SaaS release
Purpose: Collects covenant documents, validates submissions, tests covenant compliance, manages exceptions, and issues portfolio-level alerts. Moody's Lending Suite Monitoring

2. nCino Continuous Credit Monitoring with Analyst Digital Partner current cloud SaaS release
Purpose: Monitors borrower risk trends, covenant performance, and historical credit data; supports portfolio-manager review using the bank's approved lending data and workflows. nCino Digital Partners

Procedure: Commercial Loan Documentation and Covenant Compliance Monitoring

1. At loan approval, the credit team records each financial and non-financial covenant, definition, testing frequency, threshold, grace period, and waiver authority in the selected monitoring platform.
2. The bank links each covenant to the executed loan agreement, relevant amendment, borrower entity, facility, and approved credit memorandum.
3. The platform sends scheduled requests for compliance certificates, quarterly financial statements, and SEC filings through the approved borrower-document channel.

4. Moody's Lending Suite Monitoring or nCino Continuous Credit Monitoring extracts submitted data and applies only bank-approved covenant calculation rules.
5. The platform flags missing reports, approaching deadlines, low-confidence document extraction, inconsistent figures, covenant threshold breaches, and declining risk trends.
6. A portfolio manager verifies every alert against the executed agreement, the latest amendment, and the original borrower submission before treating it as a breach.
7. Confirmed breaches are routed to Credit Risk and Legal. They determine whether to issue a reservation-of-rights notice, waiver, cure-period notice, revised risk rating, or escalation to the credit-approval committee.
8. The bank retains the source documents, extracted values, calculations, reviewer decision, exception rationale, notice, waiver, and final outcome in an auditable loan record.
9. The pilot is measured by covenant-report collection timeliness, alert precision, breach-detection time, false-alert rate, number of accounts monitored per portfolio manager, and unresolved-covenant exceptions.

- Required human review and escalation: Portfolio Manager must verify all covenant breach alerts before issuing formal default or waiver notices to borrowers.
- Possible success metrics:
 - Time to detect covenant defaults (reduced from weeks to real-time).
 - *Portfolio* manager tracking capacity (accounts per FTE).
 - *Early-stage* loan loss provision reductions.
- Research evidence needed before making a recommendation:
 - *Sourced Fact*: Accenture analysis of 78 large banks found that AI leaders in risk decisioning boosted return on equity (ROE) by 125 basis points and reduced cost-to-income ratios by 452 basis points.
 - *Provisional Assumption*: Automated covenant monitoring will catch 95% of technical financial ratio breaches within 24 hours of quarterly filing ingestion.

4.3. Risk Management, Compliance, and AML/KYC Workflow 3.1: Sub-function: Commercial & Corporate Lending

- Workflow: Extracting financial statements, tax returns, and balance sheets from commercial loan applicants, spreading financial ratios, and drafting comprehensive credit risk memos.
- Current problem: Commercial credit analysts spend 15–20 hours per loan applicant manually re-keying audited financial statements and drafting repetitive credit narrative memos, delaying loan approval turnaround.
- Task frequency and complexity: Weekly per deal pipeline; High analytical and financial complexity.
- Data required and sensitivity: High Sensitivity. Corporate financial statements, tax returns, debt schedules, bank statements, industry benchmarks.
- Existing systems and integration needs: Commercial Loan Origination Systems (nCino, Moody's CreditLens), Accounting Software APIs, Financial Spreading Software, Enterprise Data Warehouses.
- Risks:
 - *Copyright*: Low risk.
 - *Privacy*: Medium-to-High risk (confidential corporate tax filings and unannounced financial results).
 - *Bias*: Medium risk (historical lending bias embedded in automated credit risk scoring).
 - *Misleading claims*: High risk (AI hallucinating financial ratios, miscalculating debt service coverage, or understating credit risk).
 - *Brand safety*: medium risk.
 - *Compliance*: Very High risk (Equal Credit Opportunity Act / ECOA, Fair Lending disclosures, OCC credit risk guidelines).
 - *Financial risk*: Extremely High (approving bad loans due to faulty AI financial ratio analysis leading to loan default).

1. Possible AI tool type:

A). nCino Automated Spreading current cloud SaaS release Purpose: Extracts, maps, reconciles, and analyzes data from tax returns, audited statements, 10-Ks, 10-Qs, and company-prepared financial statements within the commercial-loan workflow. nCino Automated Spreading

b). Moody's Lending Suite Loan Origination current cloud SaaS release

Purpose: Automates financial spreading and scoring, combines credit-risk analysis with data, and generates a first-draft credit memo for underwriter review. Moody's Loan Origination

Procedure: Commercial and Corporate Lending

1. The relationship manager submits the borrower's financial statements, tax returns, debt schedules, bank statements, and loan request through the commercial loan-origination system.
2. The bank pilots either nCino Automated Spreading or Moody's Lending Suite to extract, classify, map, and spread the financial data into approved bank templates.
3. The system calculates ratios only through bank-approved calculation logic, including debt-service coverage, leverage, liquidity, profitability, and covenant measures.
4. The system flags missing figures, conflicting values, unsupported document formats,
5. low-confidence extractions, unusual trends, and calculation exceptions for analyst review.
6. The AI tool generates a first-draft credit memorandum using verified spread data, the approved credit-policy manual, borrower information, and documented risk factors.
7. The credit analyst reconciles every extracted value, ratio, and narrative claim against the original financial statements and supporting documents.
8. The senior credit officer reviews the completed memorandum, approves or rejects the proposed credit structure, and records the final rationale. The tool does not make the lending decision.
9. The pilot proceeds to wider deployment only if it reduces memo-preparation time and manual re-keying while maintaining calculation accuracy, audit-trail completeness, fair-lending controls, and credit-quality outcomes

- Required human review and escalation: Chief Compliance Officer and Legal Department sign-off required for all policy gaps identified and action plans proposed.

- Possible success metrics:

- Time to analyze new regulatory filings (reduced from weeks to hours).
- Accuracy rate in regulatory requirement mapping.

- *Audit* deficiency reduction in regulatory compliance reviews.
- Research evidence needed before making a recommendation:
- *Sourced Fact*: KPMG survey data reveals 83% of corporate finance and compliance teams utilize AI for planning and predictive monitoring, though governance frameworks remain critical.
- *Provisional Assumption*: Semantic vector mapping between regulatory feeds and internal GRC databases will identify policy gaps with >90% recall accuracy

Workflow 3.2: Regulatory Change Management & Policy Gap Mapping

- Sub-function: Risk Management, Compliance, and AML/KYC
- Workflow: Ingesting newly published regulatory rules (SEC, OCC, CFPB, Basel III/IV) and mapping requirements against internal bank policies and operating procedures.
- Current problem: Compliance officers manually review thousands of pages of new regulatory guidance annually, risking missed obligations or delayed policy updates.
- Task frequency and complexity: Continuous / As rules are published; High legal and structural complexity.
- Data required and sensitivity: High Sensitivity. Internal policy manuals, standard operating procedures, unreleased business unit guidelines, regulatory texts.
- Existing systems and integration needs: Governance, Risk, and Compliance (GRC) platforms (MetricStream, ServiceNow GRC, IBM OpenPages), Policy Repositories.
- Risks:
 - *Copyright*: Low risk.
 - *Privacy*: Low risk.
 - *Bias*: Low risk.
 - *Misleading claims*: High risk (AI failing to identify subtle legal requirements or misinterpreting statutory intent).
 - *Brand safety*: Low risk.
 - *Compliance*: Extremely High risk (regulatory audit failure, public enforcement actions).
 - *Financial risk*: High (regulatory fines and mandatory remediation consultant costs).

Possible AI tool type □ IBM OpenPages Regulatory Compliance Management Version 9.2

Purpose: Ingests regulatory feeds, breaks rules into requirements, maps requirements to internal policies, procedures, controls, and owners, and creates traceable remediation tasks. IBM OpenPages RCM, OpenPages 9.2

- ServiceNow Integrated Risk Management, Now Assist for IRM Version 22.4.0 Purpose: Summarizes regulatory alerts, recommends potentially affected citations, controls, and policies, and supports regulatory-impact analysis and action-plan workflows. ServiceNow Now Assist for RCM, version 22.4.0 release notes
- Required human review and escalation: Chief Compliance Officer and Legal Department sign-off required for all policy gaps identified and action plans proposed.
- Possible success metrics:
 - *Time* to analyze new regulatory filings (reduced from weeks to hours).
 - *Accuracy* rate in regulatory requirement mapping.
 - *Audit* deficiency reduction in regulatory compliance reviews.
- Research evidence needed before making a recommendation:

Procedure: Regulatory Change Management and Policy Gap Mapping

1. The bank ingests approved regulatory feeds from bodies such as the SEC, OCC, CFPB, Basel Committee, and relevant national regulators into the selected GRC platform.
2. The platform records the regulatory source, publication date, effective date, jurisdiction, previous version, and original citation for every new or amended rule.
3. IBM OpenPages or ServiceNow IRM decomposes the change into individual regulatory requirements and proposes links to potentially affected internal policies, procedures, controls, business units, and previous remediation records.
4. A compliance analyst reviews every proposed mapping against the original regulatory text, confirms whether the requirement applies to the bank, and identifies missing, outdated, or conflicting internal policies.

5. The analyst creates a documented policy-gap record containing the exact regulatory citation, affected policy or control, gap description, accountable business owner, remediation action, and required completion date.
6. The responsible policy owner updates the internal policy or operating procedure. Legal and Compliance review material interpretations, high-risk gaps, and changes affecting customer treatment, AML/KYC, lending, or reporting obligations.
7. The Chief Compliance Officer approves final material policy changes. The approved policy version, implementation evidence, training records, and control-test results are retained in the GRC platform for audit purposes.
8. The pilot is evaluated using: time from regulatory publication to applicability decision, mapping precision and recall against expert-reviewed cases, percentage of remediation tasks completed by deadline, and audit deficiencies linked to regulatory change
 - *Sourced Fact*: KPMG survey data reveals 83% of corporate finance and compliance teams utilize AI for planning and predictive monitoring, though governance frameworks remain critical.
 - *Provisional Assumption*: Semantic vector mapping between regulatory feeds and internal GRC databases will identify policy gaps with >90% recall accuracy.

V. CONCLUSION

The multi-billion dollar "GenAI Divide" is not an indicator of model insufficiency, but an architectural failure. By shifting the deployment paradigm from passive, ungrounded prompt tools to deterministic, System-of-Record integrated workflows, enterprises can transition from the 95% pilot failure rate to the 5% scaled success tier. The 5-Layer Agentic Architecture presented in this paper provides a robust, governance-first blueprint that enables high-velocity media teams and highly regulated banking institutions alike to achieve measurable, risk-managed P&L impact.

REFERENCES

- [1] Bank for International Settlements (BIS), *Artificial Intelligence and Machine Learning in Financial Services: Risk Management and Compliance*

Frameworks, BIS Risk Papers, no. 112, 2024.

- [2] Board of Governors of the Federal Reserve System and Office of the Comptroller of the Currency (OCC), *Interagency Guidance on Third-Party Relationships: Risk Management*, 2023.
- [3] Gartner, *Top Strategic Technology Trends for 2025: AI Governance and Agentic Workflows*, Gartner Research Report, Doc. ID G0079812, 2025.
- [4] McKinsey & Company, *The Economic Potential of Generative AI: The Next Productivity Frontier in Financial Services and Marketing*, McKinsey Global Institute, 2025.
- [5] MIT NANDA, *State of AI in Business 2025: Bridging the Enterprise Pilot Failure Gap*, MIT Center for Collective Intelligence/NANDA Research Initiative, 2025.
- [6] National Institute of Standards and Technology (NIST), *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, U.S. Department of Commerce, 2023.
- [7] Securities and Exchange Commission (SEC), *Regulatory Notice on Algorithmic Bias, Predictive Analytics, and Customer Data Privacy in Financial Advisory*, SEC Staff Report, 2024.