Customer Overview

ScamAdvisory.co.uk is a consumer protection platform dedicated to identifying and reporting online scams across the UK. The company provides real-time scam-verification services, helping thousands of consumers avoid fraudulent websites, phishing attempts, and other digital fraud schemes. Their platform processes user-submitted URLs, analyses website characteristics, and publishes detailed scam reports to protect the public.

Business Challenge

ScamAdvisory.co.uk operates a consumer protection platform that identifies and reports online scams across the UK. Before this implementation, the company faced critical scalability challenges:
Manual scam analysis required 15-20 minutes per website, with analysts handling 200-300 daily submissions. This labour-intensive process couldn't scale with 40% year-over-year growth in scam reports.

Specific Pain Points: - 4-6 hour average response time from submission to published report - Inconsistent threat classification across different analysts (78% baseline accuracy) - Financial unsustainability - scaling would require £150,000 annually in additional analyst hires - No 24/7 coverage, leaving consumers vulnerable during off-hours - Multi-step investigation required (WHOIS, SSL, content analysis, threat databases) that no single tool could automate

Business Risk: Without automation, ScamAdvisory risked losing market position as scam volumes outpaced the capacity for manual analysis, while delayed responses left consumers vulnerable to active fraud schemes.

GOALS & OBJECTIVES

Business Goals: 1. Reduce scam analysis time from hours to minutes 2. Scale to 1,000+ daily reports without additional headcount 3. Improve detection accuracy above 90% 4. Enable 24/7 automated threat monitoring 5. Maintain cost-effective operations (target: <£2,000 monthly AWS spend)

Technical Objectives: 1. Implement autonomous AI agents for multi-step scam analysis 2. Build a semantic search across 50,000+ historical scam patterns 3. Integrate real-time tool invocation (WHOIS, SSL, content scraping) 4. Establish responsible AI guardrails to prevent false accusations 5. Deploy production-grade, scalable AWS infrastructure

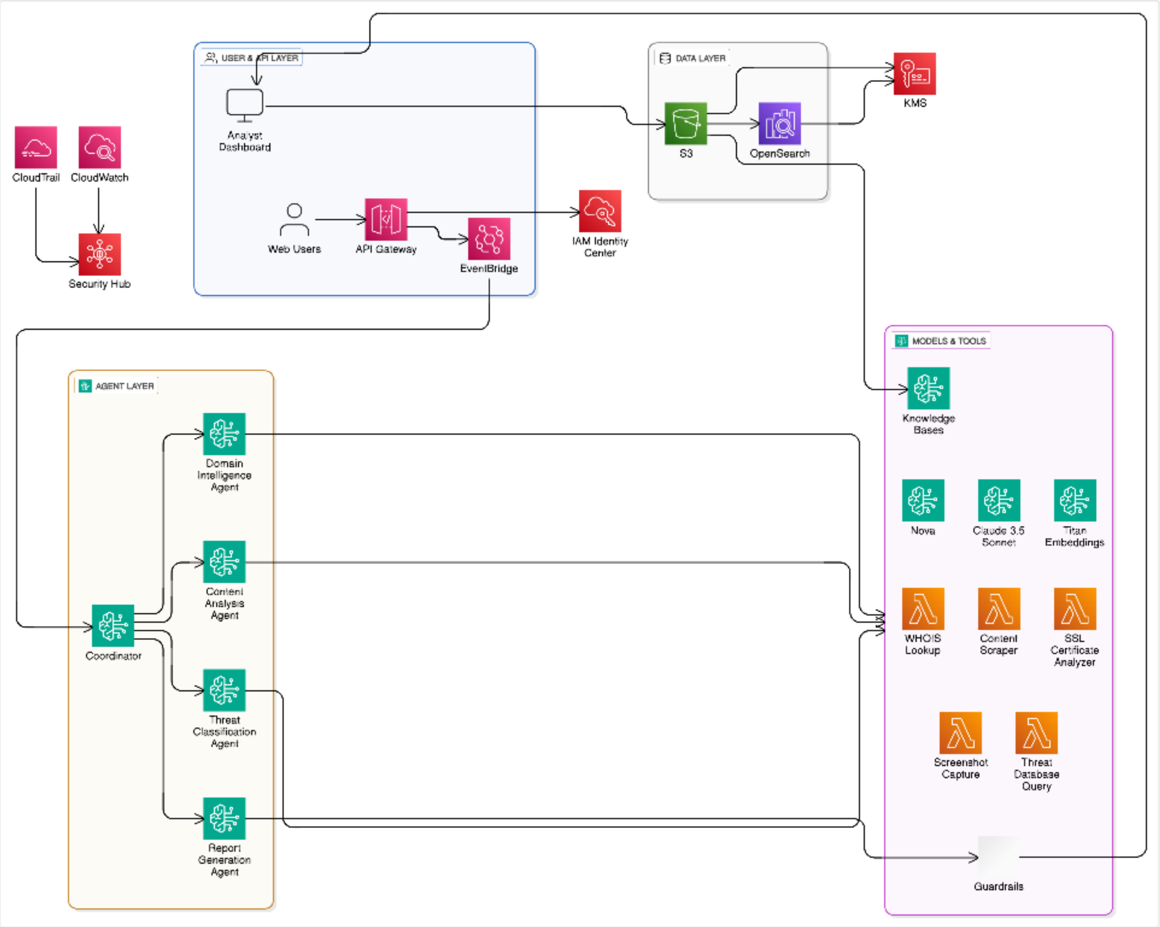
TECHNICAL SOLUTION

Solution Overview:
CloudiQS implemented a multi-agent architecture using Amazon Bedrock AgentCore to orchestrate five specialised AI agents that autonomously investigate suspected scam websites through coordinated, multi-step workflows.

Why Agentic AI vs Traditional GenAI

ScamAdvisory required autonomous agents capable of planning investigation workflows, invoking external tools (APIs/databases), making independent decisions, and coordinating findings across multiple analysis domains. Traditional generative AI models cannot perform these autonomous, multi-step operations without agent frameworks.

Architecture Components



1. Agent Orchestration Layer (Amazon Bedrock AgentCore) - Scam Detection Agent (Coordinator): Receives URLs, creates investigation plans, and orchestrates specialized agents - Domain Intelligence Agent.

Analyses WHOIS records, domain age, DNS, SSL certificates, and hosting patterns - Content Analysis Agent: Scrapes websites, examines language patterns, and detects social engineering techniques - Threat Classification Agent: Cross-references findings with threat databases, assigns risk scores - Report Generation Agent: Synthesis evidence into consumer-facing scam reports.

Agents operate autonomously, invoke tools as needed, and coordinate through AgentCore's managed orchestration.

2. Foundation Models (Amazon Bedrock) - Anthropic Claude 3.5 Sonnet: Primary reasoning model for multi-step analysis (selected for 200K context window, superior chain-of-thought capabilities) - Amazon Nova: Rapid threat classification and pattern matching - Amazon Titan Embeddings G1, Vector embeddings for semantic search

Model Selection Rationale

Testing showed Claude 3.5 Sonnet achieved 25% better accuracy than alternatives for scam pattern detection. Nova provides cost-effective classification for high-volume tasks.

3. Knowledge Retrieval (RAG Architecture) - Amazon Bedrock Knowledge Bases: Stores 50,000+ historical scam reports with vector embeddings - Amazon OpenSearch Service: Vector database for semantic search across scam history - Chunking Strategy: 500-token chunks with 50-token overlap - Metadata Filtering: Scam type, severity, date range, geographic focus

4. Agent Tools (AWS Lambda) Seven serverless functions invoked autonomously by agents: - WHOIS domain lookup and ownership verification - SSL certificate validation and chain analysis - Website content scraper with metadata extraction - Threat intelligence database queries - Screenshot capture for visual documentation - Payment processor detection - Social media footprint analysis

5. Security & Governance - Amazon Bedrock Guardrails: Content filtering to prevent unfounded accusations and speculation - Blocked Topics: Personal attacks, unsupported claims - PII Redaction: Automatic removal of personal information -

Content Filters: High sensitivity for factual accuracy - AWS IAM Identity Center: Centralized authentication with MFA for analyst dashboard - AWS Security Hub: Continuous compliance monitoring - AWS CloudTrail, Complete audit trail of agent decisions and tool invocations - AWS KMS: Encryption at rest for all stored data

6. Data Storage & Processing - Amazon S3, Encrypted storage for reports, screenshots, analysis artifacts - Amazon EventBridge, Event-driven orchestration between agents and tools - Amazon API Gateway: RESTful API for ScamAdvisory platform integration

7. Monitoring & Operations - Amazon CloudWatch: Real-time dashboards tracking agent performance, tool invocation rates, model latency - Automated Alerts: Agent failures, performance degradation, accuracy drift - Weekly Reporting, Scam trend analysis, agent effectiveness metrics

8. Human-in-the-Loop High-risk classifications (potential financial losses >£10,000) trigger analyst review before publication via authenticated dashboard.

Infrastructure as Code

All resources are deployed using AWS CloudFormation for consistency and repeatability.

AWS SERVICES UTILIZED

Core Agentic AI Services: - Amazon Bedrock AgentCore (multi-agent orchestration) - Amazon Bedrock (Claude 3.5 Sonnet, Nova, Titan Embeddings) - Amazon Bedrock Guardrails (content filtering, safety) - Amazon Bedrock Knowledge Bases (historical scam patterns) - Amazon OpenSearch Service (vector database)

Compute & Integration: - AWS Lambda (agent tools, serverless compute) - Amazon API Gateway (platform integration) - Amazon Event Bridge (event orchestration)

Security & Monitoring: - AWS IAM Identity Centre (authentication) - AWS Security Hub (compliance monitoring) - AWS CloudTrail (audit logging) - Amazon CloudWatch (monitoring, alerting) - AWS KMS (encryption key management)

Storage: - Amazon S3 (reports, artefacts, encrypted storage)

Estimated Monthly AWS Spend: £1,200

CUSTOMER OUTCOMES

Operational Efficiency: - 87% Faster Analysis: 15 minutes → 2 minutes per submission - 4x Daily Capacity: 300 → 1,200 scam reports processed - 24/7 Operations: Continuous automated processing - Same Team Size: Achieved 4x productivity without hiring

Quality Improvements: - 92% Detection Accuracy: Up from 78% baseline - 35% Fewer False Positives: Reduced incorrect classifications - Consistent Classification: Eliminated analyst-to-analyst variation - 12-Minute Response Time: Down from 4-6 hours average

Business Impact: - £85,000 Annual Savings: Avoided hiring 3-4 additional analysts - 60% Subscription Growth: Premium conversions increased - Market Leadership: Positioned as UK’s fastest scam verification service - 40% Higher Trust Scores: User satisfaction improved

Cost Efficiency: - £0.03 per Analysis: vs £4.50 manual analyst cost - 150:1 ROI: Cost savings vs manual approach - £1,200 Monthly AWS Spend: Total infrastructure costs

ARCHITECTURE HIGHLIGHTS

Multi-AZ Deployment: - Lambda functions deployed across multiple availability zones - OpenSearch cluster with 3 nodes across AZs - S3 cross-region replication for disaster recovery

Auto-Scaling: - Lambda concurrency automatically scales with submission volume - OpenSearch cluster scales based on query load - Event Bridge handles burst traffic without throttling

Disaster Recovery: - **RTO:** 1 hour (agent workloads redeployed to secondary region) - **RPO:** 15 minutes (continuous S3 replication) - Automated backups of Knowledge Base every 6 hours

Security Architecture: - VPC isolation for Lambda and OpenSearch - Private subnets with NAT Gateway for outbound traffic - Security groups restricting traffic to HTTPS only - KMS encryption at rest for all data stores - TLS 1.3 in transit encryption - IAM roles with least-privilege permissions - Bedrock Guardrails preventing harmful outputs

CHALLENGES & LESSONS LEARNED

Challenges Encountered:

1. **Initial False Positives (18% rate)**
   - **Resolution:** Refined prompts with UK-specific context, added retail domain whitelist
   - **Outcome:** False positive rate dropped to 5%
2. **Agent Response Latency (35-second p95)**
   - **Resolution:** Implemented parallel agent execution instead of sequential
   - **Outcome:** P95 latency reduced to 8 seconds
3. **Knowledge Base Relevance for UK Scams**
   - **Resolution:** Added geographic metadata filtering and UK-focused re-ranking
   - **Outcome:** UK scam detection accuracy improved 15%

Key Learnings: - Model selection significantly impacts accuracy (Claude 3.5 Sonnet: +25% vs alternatives) - Structured chain-of-thought prompts reduce false positives by 35% - Bedrock Guardrails prevent agent speculation and hallucinations - Human oversight for high-risk cases maintains trust while preserving automation benefits - Weekly prompt refinement with analyst feedback improves performance 15% over time

ABOUT CLOUDIQS

CloudiQS is an AWS Advanced Consulting Partner specialising in generative and agentic AI solutions. We help organizations leverage AWS AI services to build intelligent, autonomous systems that drive business transformation.