AHEADMG STRATEGIC OPPORTUNITY · MARCH 2026

Verity — the platform behind your *AI governance practice*

AI TEVV (Test, Evaluation, Verification and Validation) is projected to be a $10–15 billion market by 2028. AheadMG has the trust, sector depth, and client relationships to lead. This document explains the opportunity, the platform, and how the partnership works.

$10–15bn
projected AI governance market by 2028, growing 30%+ annually

2,000+
high-risk AI systems in scope across UK FS firms by end of 2026

Aug 2026
EU AI Act high-risk requirements enter full effect

Prepared by: FEAW Services Ltd · Prepared for: AheadMG Limited — Confidential

Date: March 2026 · Purpose: Strategic opportunity exploration — not a binding commercial proposal

> *"You know, I received a message this week offering help with our TEVV from a different test company. The thing is — trust is everything in AI Governance. They didn't have it, but if AheadMG starts offering TEVV we can talk."*
>
> — Director of AI & Data Governance, Financial Services (LinkedIn comment on an AheadMG post, March 2026)

This is not a hypothetical market. It is a named professional, in AheadMG's sector, with budget authority and a live need, telling you directly that your brand has the trust required to win this work — and asking you to start.

This document explains what AI TEVV is, why AheadMG is naturally positioned to deliver it, what a platform to support that practice would look like in practice, and how FEAW would build it. Commercial terms are deliberately kept to a minimum here: the goal is shared understanding of the opportunity first.

THE OPPORTUNITY

Why this is happening now — and why AheadMG

1.1 — Why this is happening now

Regulated financial services firms are deploying AI at a pace their governance frameworks cannot match. Fraud detection, credit scoring, client suitability assessment, document processing, chatbots, regulatory reporting automation — the list grows every quarter. The executive teams commissioning these systems know, at some level, that they are not being properly tested. Their governance, risk, and compliance functions know it more acutely.

The reason this has remained an unresolved problem for so long is simple: the people who understand AI well enough to test it are not the people who run QA and governance programmes, and the people who run QA programmes do not yet have the tools or frameworks adapted to AI's specific challenges. That gap is now being forced shut by regulatory deadlines.

REGULATORY DRIVER	WHAT IT NOW REQUIRES
EU AI Act (2024–2026 rollout)	High-risk AI systems must be tested before deployment and monitored on an ongoing basis. Technical documentation must prove conformity. Non-compliance carries fines up to €30m or 6% of global turnover.
FCA Model Risk (PS7/24)	Material models — explicitly including AI — require documented, independent validation. Firms must be able to explain model decisions and demonstrate ongoing monitoring.
NIST AI RMF	A voluntary but increasingly adopted framework that defines how to Govern, Map, Measure and Manage AI risk. UK firms engaging with US counterparts or investors are expected to be familiar with it.
ISO/IEC 42001	The international management system standard for AI. Emerging as the audit benchmark for firms seeking to demonstrate responsible AI governance to clients and regulators.

The EU AI Act's requirements for high-risk systems enter full effect in August 2026. FCA PS7/24 implementation timelines are now live. Firms that have not started

> building their AI governance and TEVV capability now will be scrambling by the end of 2026. This is the window — and it will not be open for long.

1.2 — Why AheadMG

The LinkedIn comment at the top of this document captures something that cannot be manufactured: a named client in the sector saying the AheadMG brand carries enough trust to have this conversation where others do not.

That trust is not accidental. It is the product of years of embedded delivery inside regulated firms. AheadMG's positioning — led by former COOs, CTOs, and QA practitioners, FSQS-registered, with a proven track record inside wealth management, insurance, and financial services — is exactly the profile a Chief Risk Officer or Head of AI Governance wants to see when commissioning something as consequential as AI validation.

Deep sector understanding 1

Years of embedded delivery inside FS, insurance, and wealth management. The regulatory context is known territory.

QA and test methodology credibility 2

AheadMG.Lens, proven test strategy and management capability, and a team with ISTQB-qualified practitioners.

Trust at the risk and governance level 3

FSQS registration, direct relationships with COO, CTO, and CRO-level stakeholders across the client base.

Independence and objectivity 4

Not a vendor of AI systems. AheadMG has no commercial interest in any particular model or platform passing validation.

5

Embedded delivery, not arms-length advice

The AheadMG model — sitting on the client's side of the table — is exactly what TEVV requires. This is not a report-and-leave engagement.

THE FRAMEWORK

What is TEVV, plainly stated

TEVV stands for Test, Evaluation, Verification and Validation. It originated in defence and aerospace software engineering — sectors where a software failure is measured in lives. NIST adopted it as the canonical framework for assessing AI systems.

For four decades, TEVV was a specialised discipline confined to defence, aerospace, and safety-critical engineering. Then AI arrived in the mainstream — and suddenly every financial services firm, insurer, and healthcare provider found themselves operating software whose behaviour they could not fully predict, explain, or guarantee. The US National Institute of Standards and Technology (NIST) formally adopted TEVV as the evaluation backbone of its AI Risk Management Framework in 2023. The EU AI Act and FCA model risk guidance followed with their own validation requirements. TEVV moved from defence laboratories to the boardrooms of regulated industry.

Testing

Run the system and check its outputs

Feed the model 500 historical client profiles. Check whether its suitability recommendations match what a human adviser would have recommended.

Evaluation

Score how well it performs across multiple dimensions

Does it perform equally well for clients over 70 as for clients under 40? What happens when data fields are missing?

Verification

Confirm it does what its specification says

The spec says the model must not recommend high-risk products to clients with a low-risk appetite score. Verify it actually enforces this, every time.

Validation

Confirm it is fit for real-world use

The model passed all tests. Does it still behave correctly when deployed against live customer data, under real production conditions?

2.2 — Why AI makes this harder

CHALLENGE	WHAT IT MEANS	WHY IT MATTERS
Non-determinism	The same input can produce different outputs on different runs	Tests must be repeated, statistical, and tracked over time.
Model drift	Behaviour changes over time as data distributions shift	A model validated in January may produce different decisions in October.

CHALLENGE	WHAT IT MEANS	WHY IT MATTERS
Invisible bias	Discrimination may be absent in aggregate but present in specific cohorts	A model with 94% accuracy overall may systematically disadvantage older customers.
Explainability	Often impossible to fully trace why a specific decision was made	FCA and EU AI Act both require firms to explain adverse AI decisions.
Training data opacity	Data the model was trained on may be too vast to audit fully	You cannot directly verify what the model has learned.

> **Because AI systems are non-deterministic, drift-prone, and often opaque, AI governance cannot be a point-in-time activity.** It must be a continuous operational process — before deployment, through deployment, and across the entire operational life of the system. This is precisely the gap that neither firms nor their existing tooling currently fills — and where AheadMG has a natural role.

2.3 — What TEVV looks like today without a platform

1. A consultant or internal analyst is assigned to “validate” an AI system before go-live
2. They review the vendor’s documentation (if it exists), run some manual test cases, and write a report
3. The report is filed in a SharePoint folder and rarely revisited
4. There is no tracking of whether the system has changed, whether the validation is still current, or whether the findings were addressed
5. When an auditor or regulator asks for evidence, the firm searches for the report, finds it is 18 months old, and scrambles to produce something current

6 Meanwhile, the system has been updated three times and is materially different from what was originally validated

This is not a process failing. It is a tooling failing. The people responsible for AI governance are intelligent and motivated. They simply do not have a system designed to support what they are being asked to do. That is what Verity provides.

THE PLATFORM

What Verity is — and what it is not

VERITY IS	VERITY IS NOT
A structured workflow platform for managing the TEVV lifecycle	A fully automated testing tool that replaces human judgement
A system of record — persistent, auditable history of every validation	A pure UI for manually entering data with no intelligence
AI-assisted tool that helps practitioners produce better test cases faster	An AI system that validates other AI systems autonomously
Generates regulatory-ready evidence documentation automatically	A compliance reporting tool that works independently of actual testing
Continuous monitoring platform tracking AI systems across their life	A one-time audit tool deployed for a single engagement

The closest analogy is Jira for AI governance: a persistent, structured system that tracks every AI system in the estate, supports the practitioners responsible for them,

and generates the evidence needed to satisfy regulators and boards — with AI assistance built into each step of the process.

3.2 — Five-stage workflow

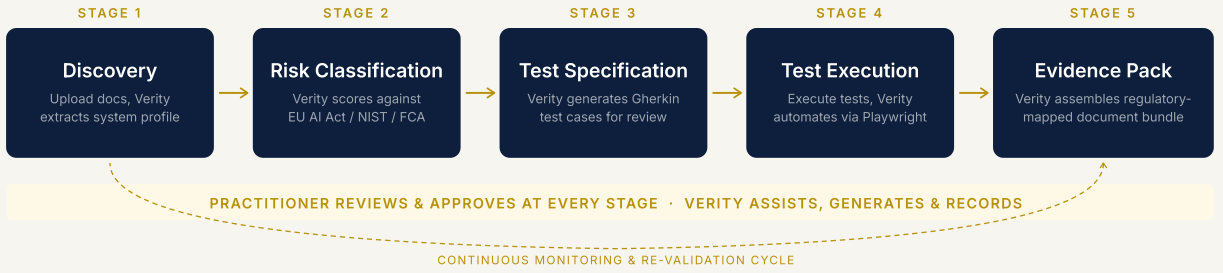


FIG 1 — The five-stage Verity TEVV workflow. The practitioner retains control at every stage; Verity assists, generates, and records.

#	STAGE	THE PRACTITIONER DOES...	VERITY DOES...
1	**Discovery**	Uploads system documentation: model cards, technical specs, API contracts, existing test reports, data dictionaries. If none exist, completes a structured intake form.	Reads all uploaded documents. Extracts system profile: purpose, inputs, outputs, risk indicators. Identifies gaps in documentation.
2	**Risk Classification**	Reviews the suggested risk classification. Confirms or overrides it. Documents rationale.	Scores the system automatically against EU AI Act risk criteria, NIST AI RMF categories, and FCA materiality indicators. Suggests a risk tier.
3	**Test Specification**	Reviews AI-generated test cases. Edits, removes, or adds cases based on professional judgement. Approves the final test specification.	Generates a structured set of TEVV test cases in plain-English Gherkin format (Given/When/Then). Covers accuracy, fairness, robustness, explainability, and safety — weighted to the system's risk tier.

#	STAGE	THE PRACTITIONER DOES...	VERITY DOES...
4	**Test Execution**	Executes tests against the AI system (manually, or via automated pipeline where available). Records results and findings. Raises defects in Jira or Azure DevOps.	Where the system exposes an API or UI, automated Playwright agents can execute test cases and record results directly. Results are captured in the platform either way.
5	**Evidence Pack**	Reviews the generated evidence pack. Adds narrative context. Approves and exports.	Assembles all inputs, classifications, test results, and findings into a structured document bundle. Maps outcomes to EU AI Act, FCA PS7/24, NIST AI RMF, and ISO 42001 requirements. Exports as PDF and Word.

3.3 — A day in the life

Scenario: A wealth manager has deployed a new AI-driven client suitability assessment tool

DAY 1 — ONBOARDING

System registration

The AheadMG consultant logs into Verity via their browser on the corporate network (no special software to install). They click ‘Register new system’. A structured form walks them through: system name, vendor, intended use, what data it takes in, what decisions it outputs. They upload the vendor’s model card and the internal deployment spec. Time: 20–30 minutes.

DAY 1 — RISK CLASSIFICATION

Automated risk scoring

Verity flags the system as High Risk under the EU AI Act: it makes recommendations affecting financial products for retail customers. The consultant reads the reasoning, agrees, and confirms the classification. The system is now formally registered in the registry with an open TEVV cycle.

DAY 2 — TEST SPECIFICATION

AI-generated test cases

Verity generates 38 Gherkin test cases. The consultant reads them. They remove 4 that are not applicable to this system, add 2 specific cases they know from experience matter for this client, and approve the specification. Time: 45–60 minutes.

DAYS 3–5 — TEST EXECUTION

Working through the test cases

The consultant works through the test cases against the live system in the client's test environment. Results are recorded in Verity as pass, fail, or observation. Two findings are raised as defects, linked directly to Jira tickets. The overall pass rate is 91% with two open remediation items.

DAY 6 — EVIDENCE PACK

Regulatory-ready documentation

The consultant clicks 'Generate Evidence Pack'. Verity produces a structured 18-page document: executive summary, risk classification rationale, test case inventory, results, findings, and a regulatory mapping table showing which EU AI Act articles and FCA PS7/24 requirements each test case addresses. The pack is exported as Word (for the client to review and sign off) and PDF (for the audit file).

ONGOING — MONITORING

Continuous lifecycle management

The system is marked Validated in the registry with two open remediation items tracked. In 90 days, the system is automatically flagged for re-validation. When the vendor releases an update next month, Verity prompts the consultant to initiate a change-triggered review cycle.

ENTERPRISE DEPLOYMENT

How this works inside a regulated corporate environment

This is the most practically important section for regulated financial services clients. The questions they will ask immediately are: where does our data go? Can it be kept off the internet? Does it require a connection to a cloud AI service? Can we use our own approved models? Verity is architected from the ground up to support all of these constraints.

OPTION 1

Cloud SaaS

Hosted by FEAW, UK datacentre

- Hosted on Azure UK South
- Multi-tenant, data isolated per client
- FEAW manages all ops & updates
- No client infrastructure needed
- **Best for:** pilot engagements

OPTION 2

Private Cloud

Client's own Azure/AWS tenant

- Deployed into client’s own cloud subscription
- Data never leaves client’s cloud boundary
- Client controls access and encryption keys
- FEAW provides deployment package & updates
- **Best for:** large FS firms with cloud governance

OPTION 3

On-Premise

Client's own data centre

- Deployed on client’s servers (Docker/Kubernetes)
- No internet connection required
- Fully air-gapped operation supported
- Updates via secure artefact transfer
- **Best for:** highly regulated / air-gapped environments

OPTION 4

Hybrid

Registry in cloud, pipeline on-prem

- Dashboard & registry in client cloud
- AI agent pipeline runs on-premise
- Sensitive data stays inside perimeter
- Evidence packs generated locally
- **Best for:** firms wanting cloud UI, internal AI

Air-gapping

Yes. This is a design requirement, not an afterthought. The on-premise deployment option supports completely air-gapped operation:

✓ The Verity platform itself runs entirely on the client's infrastructure. No calls to external services.

✓ AI agent components can use a locally-hosted AI model rather than a cloud API.

✓ Software updates are delivered as signed container images via secure transfer mechanism.

✓ Audit logs and evidence packs are stored entirely within the client's own storage infrastructure.

The trade-off in a fully air-gapped environment is that the AI model available locally will generally be less capable than a frontier cloud model. In practice, for the specific tasks Verity uses AI for — reading structured documents and generating Gherkin test cases — a capable open-weight model running locally is entirely sufficient.

AI model options

MODEL CATEGORY	EXAMPLES	WHEN TO USE
Cloud API (default for SaaS)	Anthropic Claude, OpenAI GPT-4o, Azure OpenAI	Standard deployments where internet access is permitted
Client-hosted open-weight model	Llama 3.1 70B, Mistral Large, Qwen 2.5 72B	Where data must not leave the client's environment

MODEL CATEGORY	EXAMPLES	WHEN TO USE
Client's existing approved LLM	Whatever the client has already procured	If the client already has an enterprise LLM licence
Lightweight / quantised model	Llama 3.2 8B, Phi-3 Mini	Where GPU resource is constrained

The client and AheadMG choose the model configuration that fits the client's security, performance, and budget requirements. The platform does not depend on any single AI provider.

Data classification

- No customer personal data is required by Verity
- Inputs are system documentation and synthetic/anonymised test datasets
- In private cloud or on-premise deployment, all data remains within the client's sovereign boundary
- Can be configured to operate within Microsoft Purview or equivalent frameworks
- Access controls are role-based

THE PARTNERSHIP

The AheadMG + FEAW model

This is not a technology sale. It is a partnership between two complementary capabilities. Neither FEAW nor AheadMG can deliver this alone — and that is precisely why the combination is defensible.

FEAW Brings

- Multi-agent AI engineering using Claude Code
- Next.js platform engineering and SaaS product experience
- NIST AI RMF, EU AI Act, and ISO 42001 framework knowledge
- On-premise / air-gap deployment architecture
- Ongoing platform development and AI model integration

AheadMG Brings

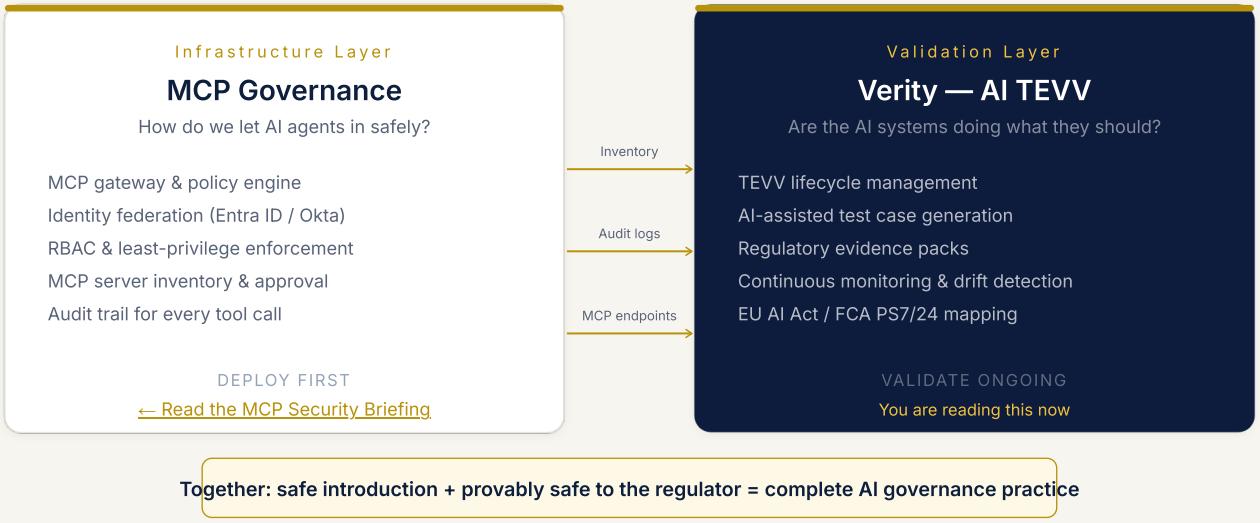
- Deep relationships with CRO, CTO, and compliance leaders in FS, insurance, and wealth management
- FSQS registration and pre-approved vendor status with financial services procurement teams
- Embedded delivery model — the consultants who would use Verity day-to-day
- QA methodology, AheadMG.Lens benchmarking capability, and test management expertise
- Client trust: the LinkedIn signal is real. When AheadMG says they offer TEVV, people listen.

> **FEAW builds and maintains the platform. AheadMG owns the client relationship and delivers the practice. Both earn from the outcome.**

THE COMPLETE PICTURE

Two halves of the same proposition

MCP governance and Verity solve different but sequential problems for the same client. Together they form a complete AI governance practice that no competitor currently offers.



MCP inventory feeds Verity Discovery. The gateway already knows every MCP server, what tools it exposes, and who's calling it. That's a pre-populated asset register for Verity to start from.

Gateway audit trail feeds Evidence Packs. Every tool call, every policy enforcement event — exactly the operational evidence that satisfies FCA PS7/24 and EU AI Act Article 9.

Verity tests MCP endpoints directly. Playwright agents can call MCP tools and assert on outputs. Nobody has built TEVV tooling specifically for agentic AI-over-MCP systems yet. That's a genuine market gap.

Data classification aligns perfectly. Verity works with system documentation and synthetic data. The MCP governance layer enforces those data boundaries at the

protocol level. Same story to a CISO.

What happens next

The purpose of this document is to open a conversation, not close a deal. The logical next step is a working session between FEAW and AheadMG to explore three questions:

1. Is there appetite to develop a TEVV practice within AheadMG, and on what timeline?
2. Which existing AheadMG client relationships represent the most natural first pilot engagement?
3. What does the right commercial and operational structure look like between FEAW and AheadMG?

FEAW is not proposing a fixed commercial structure here. The right model — whether that is a white-label licence, a co-delivery arrangement, a joint venture, or something else — should emerge from the conversation rather than be imposed before it begins.

Rob — FEAW Services Ltd · enquiries@feaw.co.uk

Related briefing

MCP Security Briefing — the threat landscape in detail

Sources

NIST (2023). *AI Risk Management Framework (AI RMF 1.0).* nist.gov

European Parliament (2024). *EU AI Act — Regulation (EU) 2024/1689.* eur-lex.europa.eu

FCA (2024). *PS7/24 — Model risk management principles.* fca.org.uk

ISO (2023). *ISO/IEC 42001:2023 — AI Management System.* iso.org

Gartner (2025). *AI Governance market projections.* gartner.com

FSQS (Financial Services Qualification System). hellios.com

Prepared by FEAW Services Ltd for AheadMG. March 2026. Confidential — not for distribution.