26 May 2026

**Office of the Chief Officer Supervision**
Tel: (+356) 21441155

**Dear Chief Executive Officer**

**Artificial Intelligence (AI) – Governance, Risk and Prudential Expectations**

You are receiving this letter as a licence holder supervised by the Malta Financial Services Authority (the "MFSA" or the "Authority") in light of the increasing adoption of Artificial Intelligence ("AI") across financial services.

**1.0 Background**

The Authority notes the introduction of the EU Artificial Intelligence Act, which establishes a harmonised framework for the development and use of AI within the Union. In cooperation with the Malta Digital Innovation Authority (MDIA), which has been designated as the market surveillance authority under the AI Act, the MFSA will continue to engage on matters relating to prudential risk, governance, and financial stability.

The MFSA continues to observe a gradual increase in the use of AI across the financial sector. While current levels of adoption among Maltese licence holders remain relatively limited, international developments indicate that the scale, complexity and criticality of such systems are expected to grow rapidly over the coming years.

AI is increasingly being integrated into core operational and decision-making processes, including risk management, customer interaction, financial crime monitoring, and internal analytics. In this context, the Authority considers it essential to ensure that such adoption is underpinned by sound governance, effective risk management, and appropriate oversight.

The Authority emphasises that the use of AI does not alter the fundamental objectives of financial regulation, namely the protection of consumers, the safeguarding of financial stability, and the preservation of market integrity. However, AI has the potential to amplify

existing risks and introduce new vulnerabilities, particularly where systems operate with a degree of autonomy, opacity, or reliance on external providers.

**2.0 Scope**

The Authority's primary objective in issuing this letter is to ensure that the adoption of artificial intelligence within the financial sector develops in a manner that is consistent with sound prudential principles, effective governance, and the overarching objectives of financial regulation.

While the current level of AI adoption across Maltese licence holders remains relatively limited, the Authority considers that the pace of technological development, coupled with increasing accessibility of AI tools, is likely to result in a rapid expansion of use cases in the near term. In this context, the Authority expects firms to adopt a forward-looking approach and to ensure that governance, risk management, and internal control frameworks are capable of supporting such developments in a sustainable and controlled manner.

This letter is therefore intended to achieve three key outcomes:

Firstly, it seeks to ensure that licence holders recognise AI as a prudentially relevant risk area, rather than solely as an operational or technological enhancement. Firms are expected to assess how AI may impact their risk profile, decision-making processes, and overall resilience, and to ensure that such considerations are embedded within existing governance and risk management structures.

Secondly, the Authority expects this letter to prompt a structured internal assessment within firms, particularly at Board and senior management level. Licence holders should critically evaluate their current and anticipated use of AI, identify any gaps in governance, oversight, or controls, and take appropriate remedial action where necessary. This includes ensuring that accountability is clearly defined, that third-party dependencies are understood and managed, and that sufficient expertise exists to oversee AI systems effectively.

Thirdly, the letter aims to promote greater consistency and preparedness across the market. By setting out the Authority's expectations at an early stage, firms are provided with the opportunity to align their frameworks proactively, thereby reducing the risk of fragmented or inconsistent approaches to AI adoption across the sector.

The Authority will take these expectations into account in the course of its ongoing supervisory activities. Licence holders should therefore be in a position to demonstrate, upon request, that AI-related risks have been appropriately considered and that governance and control frameworks are commensurate with both current and anticipated levels of AI integration.

Ultimately, the Authority expects that AI is deployed in a manner that enhances, rather than undermines, financial stability, market integrity, and consumer protection. Firms that fail to adequately address the risks associated with AI may be subject to increased supervisory scrutiny and, where necessary, further supervisory intervention.

**3.0 Supervisory Engagement and Market Observations**

In 2025, the Authority undertook a cross-sectoral assessment to better understand the extent of AI adoption within the Maltese financial services sector. This exercise formed part of broader supervisory efforts aimed at identifying emerging risks and ensuring that firms remain adequately prepared for evolving technological developments.

The assessment covered a broad population of licence holders and focused on the use, governance, and oversight of AI systems. The findings indicate that, while awareness of AI is increasing, its implementation remains at an early stage across most entities.

In particular, the Authority observed that a significant proportion of firms:

- do not yet have a formal or board-approved AI strategy;
- rely primarily on externally provided tools, particularly generative AI solutions;

- have limited internal expertise and governance structures in place to oversee AI deployment.

While these findings reflect a relatively low level of immediate systemic risk, they also suggest that AI adoption may accelerate without a corresponding strengthening of governance and control frameworks.

## 4.0 Key Observations

### 4.1 Governance and Oversight

The Authority observed that governance frameworks relating to AI remain uneven across the sector. In many cases, responsibility for AI systems is not clearly defined, and oversight at Board and senior management level is limited.

In the absence of formal governance structures, there is an increased risk that AI systems are deployed without sufficient challenge, validation, or alignment with the firm's risk appetite. Particular attention should be given where AI systems support, enable, or influence critical or important functions, as understood under the Digital Operational Resilience Act (DORA), given the heightened operational, prudential, and resilience implications associated with such functions. This is particularly relevant where such systems are used in processes that may impact customers, financial decisions, or market outcomes.

- **MFSA Expectations**

The Authority expects licence holders to ensure that AI adoption is supported by clear governance arrangements. Boards and senior management must retain effective oversight of AI systems and remain accountable for their use, including where such systems are provided by third parties.

Responsibilities for AI systems should be clearly assigned, and firms should ensure that sufficient expertise exists within the organisation to provide effective challenge and oversight. This includes ensuring that roles across the three lines of defence are appropriately defined

and that no single function exercises unchecked control over the design, deployment and validation of such systems.

### 4.2 Third-Party Dependencies and Concentration Risk

The Authority noted a strong reliance on third-party providers for AI capabilities, including cloud service providers, model developers, and data vendors. In many cases, firms have limited visibility over the underlying functioning of these systems.

This reliance introduces risks relating to concentration, operational dependency, and reduced control over critical processes. These risks may become more pronounced as AI adoption scales across the sector.

- **MFSA Expectations**

Licence holders are expected to treat AI-related outsourcing in line with existing outsourcing and third-party risk management frameworks. Firms must ensure that they retain sufficient control, oversight, and understanding of any externally provided AI systems.

In particular, firms should assess concentration risk and avoid excessive reliance on a limited number of providers, especially where such dependencies may impact critical operations.

### 4.3 Model Risk and Reliability

The Authority observed that AI systems, particularly those based on advanced or generative models, may produce outputs that are inconsistent, inaccurate, or difficult to interpret. In addition, such systems may evolve over time, resulting in changes to behaviour that are not immediately apparent.

Where AI is used in critical processes, these characteristics may undermine the reliability of decision-making and increase exposure to model risk.

- **MFSA Expectations**

Firms are expected to ensure that AI systems are subject to appropriate validation, testing, and ongoing monitoring. This includes the identification of potential model limitations, the implementation of controls to detect model drift, and the establishment of clear escalation mechanisms where issues arise.

The Authority expects that firms are able to understand, explain, and evidence how AI systems operate, particularly where outputs influence financial or customer outcomes. In this regard, firms should ensure that sufficient documentation, audit trails, and model governance artefacts are maintained to support effective internal challenge and supervisory review. Where firms are unable to adequately explain or evidence system behaviour, consideration should be given to restricting or redesigning the use of such systems, particularly in higher-risk contexts.

### 4.4 Data Governance

The effectiveness of AI systems is inherently dependent on the quality and integrity of the data on which they rely. The Authority observed that data governance frameworks are not always sufficiently developed to support the use of AI.

Inadequate data governance may result in biased outputs, incorrect conclusions, or regulatory breaches, particularly where personal or sensitive data is involved.

- **MFSA Expectations**

Licence holders are expected to implement robust data governance frameworks supporting AI use. This includes ensuring that data is accurate, relevant, and appropriately validated, and that data flows and usage are clearly understood and documented.

Firms should also ensure that data used in AI systems is consistent with applicable regulatory requirements and internal policies.

### 4.5 Operational and Systemic Considerations

While current levels of AI adoption are limited, the Authority notes the potential for broader systemic implications as usage increases. In particular, the use of common models, shared data sources, and third-party providers may lead to increased interconnectedness across firms.

In stress scenarios, such dynamics may contribute to correlated behaviour or amplify existing market vulnerabilities.

- **MFSA Expectations**

Firms are expected to take a forward-looking approach to AI risk, considering not only firm-level impacts but also potential system-wide implications. This includes assessing dependencies, identifying potential single points of failure, and ensuring that contingency measures are in place.

**5.0 Self-Assessment and Internal Review**

In order to support licence holders in assessing their exposure to AI-related risks, the Authority has developed a structured self-assessment framework, set out in Annex 1 to this letter.

This framework is designed to provide firms with a practical tool to:

- identify AI use cases across their operations;
- map dependencies, including third-party providers;
- assess governance, accountability, and oversight arrangements;
- evaluate the adequacy of existing risk management and control frameworks.

The Authority expects licence holders to complete this assessment in a **comprehensive and critical manner**, ensuring that it reflects both current use of AI and anticipated future developments.

At this stage, firms are not required to submit the results of this exercise to the Authority. However, licence holders should be in a position to demonstrate that:

- the assessment has been performed;
- its outcomes have been considered at Board and senior management level; and
- any identified gaps are being addressed through appropriate remedial actions.

The Authority may, as part of its ongoing supervisory activities, request evidence of this assessment and evaluate the extent to which firms have aligned their frameworks with the expectations set out in this letter.

**6.0 Conclusion and Next Steps**

This letter sets out the Authority's key observations and supervisory expectations with respect to AI adoption. In doing so, it aims to support licence holders in recognising AI as a prudentially relevant risk area, to prompt a structured internal assessment at Board and senior management level, and to drive greater consistency and preparedness across the market. The letter should therefore be read as a basis for firms to evaluate the adequacy of their existing governance, risk management, and control frameworks, and to take timely action where gaps are identified. The expectations set out in this letter are intended to complement existing supervisory, regulatory, and legislative obligations and should be read without prejudice to any applicable Acts, Regulations, Rules, sector-specific guidelines, or other regulatory requirements to which licence holders may be subject.

The Authority will continue to monitor developments in AI adoption and will integrate AI-related considerations into its ongoing supervisory activities, including onsite inspections and thematic reviews.

Particular attention will be given to:

- governance and oversight frameworks;
- third-party dependencies;
- the use of AI in critical or customer-impacting processes;

- alignment between AI adoption and firms' risk appetite.

Licence holders are expected to take proactive steps to strengthen their governance and risk management frameworks in anticipation of increased AI adoption.

The Authority also recognises that the effective oversight of AI systems requires an appropriate level of technical understanding across firms. In this regard, the Authority will, through the Financial Supervisors Academy (FSA), support licence holders by offering targeted training and capacity-building initiatives aimed at enhancing understanding of AI-related risks, governance, and supervisory expectations.

**Yours sincerely,**

**Malta Financial Services Authority**

Christopher P. Buttigieg
Digitally signed by Christopher P. Buttigieg
Date: 2026.05.26 09:30:23 +02'00'

Christopher P. Buttigieg
**Chief Officer Supervision**

Alan DECELIS (Signature)
Digitally signed by Alan DECELIS (Signature)
Date: 2026.05.26 09:40:14 +02'00'

Alan Decelis
**Head of Supervisory ICT Risk & Cybersecurity**

MFSA MALTA FINANCIAL SERVICES AUTHORITY

## Annex 1: AI-Enabled Process & Vendor Mapping Self-Assessment — Question Register

| Topic | Question | Evidence you will need | What you must produce |
|---|---|---|---|
| **A. Process & Tool Inventory** | **1) End-to-end process inventory completeness (Ops/Mgmt/Board)**<br>*Have you listed ALL business-critical processes and broken each one into its sub-steps, and for every sub-step named every tool actually used (including office automation and small scripts)?* | Annex A register; process maps or swim-lanes; screenshots or SOPs showing the tools used at each step. | Total number of processes; total number of sub-steps; total number of tools; tools-per-process (minimum/average/maximum). |
| | **2) Vendor & sub-vendor (4th-party) identification (Mgmt)**<br>*For EACH tool in use, have you identified the primary vendor AND all material sub-vendors (cloud/hosting provider, AI model provider, key data providers, and any open-source components listed in a software bill of materials)?* | Contracts, procurement due-diligence files, vendor responses, and Software Bill of Materials (SBOM) or architecture diagrams. | Number of vendors per process; percentage of tools that involve 4th parties; top-5 vendor concentration (% of all tools); CSP concentration by provider and region. |
| | **3) AI capability detection & classification (Ops/Mgmt)**<br>*For EACH tool, have you confirmed whether it includes AI (Yes/No)? If Yes, have you recorded the AI type (ML, GenAI, NLP, CV, agentic tools), who provides it (in-house or external), where it runs (on-premises or cloud and region), and what exact function it performs in that process step?* | Vendor documentation, technical product sheets, discovery testing results, and procurement questionnaires where AI capability is explicitly answered. | Total number of AI-enabled tools (overall and per process); breakdown by AI type and by provider (internal vs external). |
| | **4) Autonomy, materiality & impact tagging (Mgmt)**<br>*Is EACH AI use tagged with: autonomy level (None/Assist/HITL/Partial/Full), materiality (Low/Medium/High), and whether it can affect customers or markets?* | Model inventory, criticality criteria, approvals showing who authorised the autonomy level and safeguards. | Number of high-materiality AI uses; percentage of AI uses that are Partial/Full automation; split between customer-facing and internal uses. |
| **B. Data Governance & Flow Mapping** | **5) Data lineage & flows per AI-enabled tool (Ops/Mgmt)**<br>*For EACH AI use, have you mapped: inputs and outputs; whether personal data or special categories are involved; which datasets are used for training versus day-to-day use (inference); where the data are stored and processed; and any cross-border transfers?* | Data flow diagrams, Records of Processing Activities (RoPA), lineage maps from data governance tools. | Number of documented data flows; number using special categories; number with cross-border transfers; percentage of AI uses with a complete lineage map. |
| | **6) Lawful basis, consent, and reuse controls (Mgmt)**<br>*Is the lawful basis documented for EACH dataset used by AI? Are customer notices or consents up to date? Are there rules and approvals for reusing data across tools and processes?* | Data Protection Impact Assessments (DPIA), Legitimate Interests Assessments (LIA), privacy notices, reuse approvals. | Percentage of datasets with a recorded lawful basis and notice; number of data reuse cases; number of cases lacking consent/notice updates. |
| | **7) Bias, performance & drift monitoring coverage (Ops/Mgmt)**<br>*Do ALL AI uses have documented tests before deployment (accuracy and fairness) AND ongoing monitoring after deployment (performance, drift, and fairness thresholds) with clear triggers for remediation?* | Test protocols, fairness/accuracy reports, monitoring dashboards, issue logs. | % of AI uses with full pre-deployment tests; number without drift monitoring; number of remediations in the last 12 months. |
| **C. Governance, Accountability & Board MI** | **8a) Ownership & RACI per process/tool/vendor (Mgmt/Board)**<br>*For EACH line in the register, is there a named business owner, model owner, data owner, and vendor owner, ensuring no conflicts (e.g., the same person not both building and approving)?* | RACI matrix, policy extracts, committee Terms of Reference (ToR). | Percentage coverage with named owners; number of ownership gaps; number of conflicts of interest identified and resolved. |

| Topic | Question | Evidence you will need | What you must produce |
|---|---|---|---|
| | **8b) Cross-functional oversight (FCC, Compliance, IT involvement)**<br>*Are FCC, Compliance, and IT functions appropriately and consistently involved in the oversight of AI uses that touch their remit (e.g., FCC for KYC/TM/AML scenarios; Compliance for legal requirements and customer protections; IT for change control, security, resilience)?* | Committee minutes, attendance records, signoffs, change tickets showing FCC/Compliance/IT input. | % of AI uses where FCC/Compliance/IT participated in design/validation/ongoing review (where applicable); names of standing forums/committees and cadence. |
| | **8c) Calibration & maintenance operating model (internal / joint / external)**<br>*Does calibration and ongoing maintenance of the AI system happen internally, jointly with the vendor, or is it managed externally? Are decision rights, change approval, testing responsibilities, roll-back/kill-switch, and vendor access controls clearly defined and evidenced?* | Operating procedures, vendor contracts/addenda, change logs/CAB minutes, model cards, access reviews. | Count and % of AI uses with Internal vs Joint vs External-only calibration/maintenance; for External-only cases: list compensating controls; who can change models/prompts/thresholds and where documented. |
| | **9) Board MI completeness on AI & vendor risk (Board)**<br>*Does the Board receive regular reports that show: AI uses per process; vendor and 4th-party concentrations; incidents and customer detriment; and an overall AI risk heatmap with actions and deadlines?* | Board agendas and packs, action logs with owners and due dates. | Reporting frequency; copies of the last two packs; list of open actions and how old they are. |
| **D. Third-Party Risk, Resilience & Change** | **10) Outsourcing & ICT risk controls per vendor (Mgmt)**<br>*For ALL vendors supporting AI uses, do contracts cover SLAs, security, audit rights, incident notification, data location, exit/portability, and approval for sub-vendors?* | Contracts and addenda, due-diligence results, vendor responses. | Percentage of vendors that meet ALL key clauses; number of gaps by clause; number of high-risk vendors with gaps. |
| | **11) Fourth-party visibility & concentration risk (Mgmt)**<br>*Can you quantify how much you rely on the top providers (cloud/AI model/data)? Can you spot single points of failure or regional dependencies (e.g., all in one data centre region)?* | Aggregated Annex A register; a short concentration analysis note or slide. | Share of AI uses dependent on the top-3 providers; (optional) a simple concentration index; number of critical processes depending on a single provider. |
| | **12) Change management, rollback & kill-switch (Ops/Mgmt)**<br>*Do AI-enabled tools follow change control (versioning, peer review/approval, testing such as canary/A-B where suitable) and have a tested rollback plan and an accessible kill-switch?* | Change logs, Change Advisory Board (CAB) minutes, incident post-mortems/root-cause analyses. | Percentage of AI uses with a documented rollback plan and kill-switch; number of emergency rollbacks or drills in the last 12 months. |
| **E. Conduct, Prudential & Financial Crime** | **13) Customer-facing AI: disclosure & human escalation (Ops/Mgmt)**<br>*Are customers clearly informed when they interact with AI? Is it easy for any customer to reach a trained human, and do you track how quickly the handover happens, especially for vulnerable customers?* | Scripts, chatbot configurations, QA reports, service dashboards. | Number of channels using AI (web, app, phone); percentage with AI disclosure; percentage with tracked human escalation; average time-to-human. |
| | **14) Financial exclusion & fairness in AI-supported processes**<br>*Have you assessed whether these AI-supported decisions create risks of bias or financial exclusion? Are outcomes explainable, subject to effective challenge, and governed as sensitive or high-risk use cases with clear management accountability?* | Model and materiality documentation, fairness and bias testing reports, relevant customer communications and case logs, records demonstrating management oversight and governance. | Number of AI use cases used in segmentation, credit, pricing, and advice; number that have completed fairness or financial-exclusion assessments; how outcomes differ across key |

| Topic | Question | Evidence you will need | What you must produce |
|---|---|---|---|
| | | | customer groups; number of manual overrides, customer challenges, and complaints. |
| | **15a) Systemic interaction & feedback risks (all AI)**<br>*Have you assessed shared data/model dependencies, potential feedback loops, single points of failure, and alert/decision volatility for this AI, with scenarios and manual overrides/kill-switches defined?* | Scenario packs, playbooks, change/incident logs. | Number of scenarios run; number of overlaps in providers/data; number of dates/results of drills (e.g., kill-switch, rollback, manual handover). |
| | **15b) FC strategy for AI (where applicable)**<br>*Does the firm have a documented strategy for using AI in Financial Crime functions (AML, fraud, screening, KYC) — covering objectives, scope, roles, controls, and periodic review?* | FC AI strategy, policy addenda, board/committee approval. | Strategy document name/date; next review due; scope of AI uses covered. |
| | **15c) Alignment with FC risk assessment & risk appetite**<br>*How does AI use in FC align with your enterprise FC risk assessment and risk appetite (including thresholds for false positives/negatives, customer impact, and resourcing)?* | Business risk assessment, risk appetite statements, MI. | Reference to latest FC risk assessment; mapping notes to appetite metrics and thresholds. |
| | **15d) Missed-case learning**<br>*Has the firm identified cases that AI initially missed (e.g., confirmed SARs/STRs later found by humans, counterpart systems, or post-event reviews), and fed these back for model improvement?* | Case reviews, QA/QC logs, model change tickets. | Number of missed cases detected; percentage reviewed; number of model updates stemming from missed-case analysis. |
| | **15e) Review/recalibration thresholds**<br>*What quantitative thresholds trigger model review or recalibration (e.g., precision/recall drift, false positives/negatives, case aging, throughput, alert volatility)?* | Monitoring dashboards, threshold register, CAB minutes. | List of thresholds and current values; number of threshold breaches in last 12 months; outcomes. |
| | **15f) Use-case risk classification**<br>*Has the firm assessed whether each FC AI use-case is high-risk (from financial crime or customer impact) with proportionate controls (HITL, sampling, second-line review)?* | Risk classification criteria, approvals, control design. | Count/percentage of high-risk FC AI uses; control overlays applied. |
| | **15g) TM — Typology coverage (complete only if TM AI is used)**<br>*Which typologies is the TM model designed to detect (e.g., layering, smurfing, mule activity, sanctions evasion, fraud), and how is coverage validated?* | Typology library, validation reports, red-team tests, MLRO approval. | Typology list; validation frequency; gaps & remediation plan. |
| | **15h) TM — Feature/rule classification & risk-rating (complete only if TM AI is used)**<br>*How are data points/features and rules classified and risk-rated? Are proxies for protected characteristics controlled?* | Feature governance pack, fairness testing artifacts, MLRO approval. | Feature/rule catalog with risk weights; fairness checks. |
| | **15i) TM — Operating mode vs rules-based (complete only if TM AI is used)**<br>*Does the AI replace, complement, or work in tandem with rules-based scenarios (e.g., as a prioritizer)? Is there a challenger or back-stop?* | Design docs, Policies & Procedures, AML Manual. | Mode per scenario; presence of challenger/back-stop. |

| Topic | Question | Evidence you will need | What you must produce |
|---|---|---|---|
| | **15j) TM — Alert-volume stability monitoring (complete only if TM AI is used)**<br>*Is there monitoring for sudden drops (or surges) in alert volumes or case conversions, with investigation and escalation?* | MI dashboards, incident logs. | KPIs on alert volumes & conversion; number of anomalies and investigations. |
| | **16) Market & portfolio AI correlation assessment (if applicable) (Mgmt/Board)**<br>*Have you analysed shared data/model dependencies, correlation/herding risks, and feedback loops, and run stress scenarios with clear kill-switches or manual overrides?* | Stress-test packs, scenario results, playbooks showing who does what and when. | Number of scenarios run; overlaps in providers/data; results from a 'liquidity withdrawal' drill (how you would reduce or stop activity safely). |