CYBER STANDARDS DOCUMENT

Artificial Intelligence Risk Assessment

ABSTRACT:
This document provides a structured approach to assessing and managing security risks associated with AI adoption in law enforcement. It ensures that AI technologies are evaluated at every stage of their lifecycle, meeting security and operational requirements.

This standard adheres to the National Policing Community Security Policy Framework and is a suitable reference for community members, notably those who build and implement IT systems on behalf of national policing.

ISSUED	May 2025
PLANNED REVIEW DATE	May 2026
DISTRIBUTION	Community Security Policy Framework Members

POLICY VALIDITY STATEMENT
This standard is due for review on the date shown above. After this date, this document may become invalid.

Cyber Standard users should ensure that they are consulting the currently valid version of the documentation.

CONTENTS

VERSION: 1.0
DATE: 07/05/25
REFERENCE: PDS-CSP-STD-AIRA

DOCUMENT SIZE: 26-Page Document
CLASSIFICATION: OFFICIAL

VERSION: 1.0
DATE: 07/05/25
REFERENCE: PDS-CSP-STD-AIRA

DOCUMENT SIZE: 26-Page Document
CLASSIFICATION: OFFICIAL

Community Security Policy Commitment

National Policing and its community members recognise that threats to policing information assets present significant risk to policing operations. National Policing and its community members are committed to managing information security and risk and maintaining an appropriate response to current and emerging threats, as an enabling mechanism for policing to achieve its operational objectives whilst preserving life, property, and civil liberties.

This risk assessment in conjunction with the National Policing Community Security Policy Framework and associated documents sets out National Policing requirements for the use of Artificial Intelligence in Policing.

Introduction

As Artificial Intelligence (AI) becomes increasingly integrated into law enforcement, it is essential that policing organisations have a structured process to assess and manage the risks associated with AI technologies. While AI has the potential to significantly enhance the work of the policing community, it must be developed, deployed, and operated securely and responsibly.

Without proper oversight, AI introduces unique security risks and operational challenges that require careful consideration. To ensure AI technologies function ethically and securely, the policing community must proactively identify and mitigate these risks at every stage of AI adoption. A secure-by-design approach is essential, embedding security from the outset and enforcing it throughout the AI system's lifecycle. This approach ensures cyber resilience, making security a core requirement throughout the entire lifecycle of AI technologies.

As policing organisations begin integrating AI technologies across their operations, it is important that they apply the recommendations from this risk assessment to ensure AI is implemented securely and responsibly.

Definitions

There are many definitions available for Artificial Intelligence (AI). For the purposes of this document, we have adopted and expanded on the definition provided by the NPCC endorsed – 'Principles for Using Artificial Intelligence (AI) in Policing', written by Science & Technology in Policing. While we reference AI throughout this document, it is intended that reference to AI, covers all of the below.

What is Artificial Intelligence? [1]

There is no definitive definition of Artificial Intelligence (Alan Turing Institute, 2021), and AI is often used to refer to related applications such as automation, neural networks, and machine learning. To bring clarity for policing, we adopt the following definitions:

- **Artificial intelligence (AI)** refers to a machine that learns, generalises, or infers meaning from input, thereby reproducing or surpassing human performance. An example is using image analysis to determine whether a video contains sexual activity with a child. The term AI can also be used loosely to describe a machine's ability to perform repetitive tasks without guidance.
- **Machine learning (ML)** refers to algorithms that leverage new data to improve their ability to make predictions or decisions, without having been explicitly programmed to do so. ML is a widely used form of AI that has contributed to innovations such as speech recognition and fraud detection.
- **Advanced Data Analytics (ADA)** uses subject matter expertise and techniques that are typically beyond those of traditional business intelligence to extract insights and make recommendations from complex data. The techniques vary widely, from data visualisation to complex linear models to language analytics. An example is the use of Risk Terrain Modelling to quantify environmental factors that shape risk mapping and resource deployments.

There is other related AI terminology, which start to overlap with the above, but are included here for completeness. The above and below are considered the most common, but there are others. The additional definitions are:

- **Generative Artificial Intelligence (GAI)** is artificial intelligence capable of generating text, images, or other media, using generative models. Generative AI models learn the patterns and structure of their input training data and then generate new data that has similar characteristics.[2]
- **Large Language Models (LLMs)** are a subset of GAI, where an algorithm has been trained on a large amount of text-based data, typically scraped from the open internet, and so covers web pages and - depending on the LLM - other sources such as scientific research, books, or social media posts.[3] Examples include ChatGPT, DeepSeek, Gemini, X's Grok and Meta's LLaMA.

[1] NPCC - Principles for Using Artificial Intelligence (AI) in Policing
[2] Generative artificial intelligence - Wikipedia
[3] ChatGPT and LLMs: what's the risk - NCSC.GOV.UK

- **Natural Language Processing** is a computer's attempt to "understand" spoken or written language. It must parse vocabulary, grammar, and intent, and allow for variation in language use. The process often involves machine learning.[4]

Owner

National Chief Information Security Officer (NCISO).

Purpose

This risk assessment should empower policing to leverage artificial intelligence responsibly.

It is intended to help policing organisations meet Community Security Policy requirements in a relatively new and fast evolving technology area.

It is important that policing can be innovative with this technology, and so this risk assessment seeks to provide the guardrails, so that innovation can be carried out safely and securely and does not put policing at unnecessary risk of losing public trust and confidence through consequential data loss or loss of policing services.

Audience

Members of the Policing Community of Trust.

More specifically the standard is targeted at, architects, developers, data scientists and security experts tasked with designing and building solutions, applications and plugins leveraging AI related technologies.

The following should also be aware of the content of this standard, in order that they can provide appropriate oversight and governance of the use of AI related technologies within policing:

- Senior Information Risk Owners (SIROs)
- Information Asset Owners (IAOs)
- Information & Cyber risk practitioners and managers
- Auditors providing assurance services to PDS or policing.

[4] AI Glossary: Artificial intelligence, in so many words | Science

Finally, Policing's reliance on third parties means that suppliers acting as service providers or developing products or services for PDS or policing, should also be made aware of and comply with the content of this standard, in relation to their work on Policing systems and data.

Scope

In scope for this risk assessment are cyber security considerations and requirements for the:

- Acquisition and implementation of solutions that incorporate AI.
- Development of solutions with integrated AI.
- Use of AI technologies, e.g. ChatGPT, Gemini and LLaMA.

Not all the requirements listed below will apply in the same way across each use case. Forces should apply this assessment proportionately, identifying whether a requirement is met directly, needs to be evidenced by the supplier, or acknowledged as a residual risk to be documented and managed appropriately.

This risk assessment is intended to be used in conjunction with the NPCC AI Principles outlined in the Covenant for Using Artificial Intelligence (AI) in Policing[5] to ensure that both cyber security and ethical and legal considerations are comprehensively addressed when adopting AI technologies in policing.

Requirements

This section details the minimum requirements for the acquisition, development and use of AI to protect policing from the loss of confidentiality, integrity or availability of the data or loss of availability of the systems and services it relies upon to meet policing outcomes.

The newness of AI technologies and their potential for driving significant change suggests that these systems demand an entirely fresh set of control requirements to monitor, police and curate them. This is not necessarily the case. These systems in many ways are similar to existing technologies in use across policing organisations today, making them manageable in the same way as those familiar systems.

Most control requirements relating to the acquisition and use of AI and AI-based technologies can be found in the existing standards which have been written, or are being written to support the National Policing Community Security Policy and Principles. These guidelines apply to AI technologies in the same way they

[5] NPCC - Principles for Using Artificial Intelligence (AI) in Policing

apply to other digital solution. However, AI technologies and AI integrated technologies can present a unique set of vulnerabilities, and so the requirements below are to address these.

These requirements are not intended to replace those in other standards and while there will be duplication, they are documented here, as they are particularly pertinent to the acquisition, development, and use of AI-based technologies.

This risk assessment and these requirements do not standalone; it is important that all cyber security standards are considered during the acquisition and development of AI-based solutions.

Existing standards that should be consulted include:

- Artificial Intelligence & LLM (Large Language Models) Standard
- Security Architectural Principles
- System Management Standard
- System Access Standard
- System Development Standard
- Threat & Incident Management Standard
- Cryptography Standard
- Third Party Assurance for Policing Standard
- Technical Security Management Standard
- Application Management Standard
- Information Management Standard
- Information Assurance Standard
- Secure by Design Guideline
- Business Continuity Standard
- Network Security Standard

The above is not an exhaustive list but will provide a solid base when developing or deploying AI-based technologies.

Reference	Minimum requirement	Control reference	Compliance Metric / Artefacts
1	AI Adoption and Governance		
1.1	Before adopting an AI technology, policing organisations must assess whether AI is the most appropriate tool to meet their needs. • Evaluate if AI is necessary and how it will effectively meet your requirements. • Follow a 'Secure by Design' (SbD) methodology. • Threat modelling must be conducted at each stage of AI development, deployment and integration of new components.	**NIST CSF:** PR.IP.2 **ISO 27002:2022:** 8.27 **ISF SOGP:** SD1.2	SbD process documentation SbD project artefacts Threat Modelling Reports Secure Development Lifecycle
1.2	Policing organisations must exercise caution when using AI technologies, particularly for law enforcement purposes, to mitigate risks related to data confidentiality, integrity and authenticity while supporting responsible decision-making. A formal risk assessment must be conducted before adopting an AI technology, with formal approval from the appropriate risk owner.	**NIST CSF:** ID.AM.3&5/ PR.DS.1&2&3/ PR.IP.6/ID.GV.3 **ISO 27002:2022:** 5.1/5.12/ 5.13/5.14/5.32/5.33/ 5.34/5.9/8.1/ 8.11 **ISF SOGP:** IM1.1/1.2/1.3/1.4	Evidence that the following have been completed: • Business Impact Assessments • Data Protection Impact Assessment • Risk registers • Risk Mitigation Plans • Information Risk Management Framework

Reference	Minimum requirement	Control reference	Compliance Metric / Artefacts
1.3	A business impact assessment must be conducted at the outset to establish a foundation for evaluating both potential benefits and risks of AI adoption, ensuring alignment with the organisation's objectives and security standards.	**NIST CSF:** ID.RA.4 **ISO 27002:2022:** N/A **ISF SOGP:** IR2.2	Evidence that the following have been completed: • Business Impact Assessments • Data Protection Impact Assessment (DPIA) • Risk registers • Risk Mitigation Plans • Information Risk Management Framework
1.4	Ensure you have the requisite skilled people, process and technology elements required to set up, run and maintain operational AI technologies and cope with their outputs.	**NIST CSF:** PR.DS.4/PR.MA.1 PR.PT.4&5/PR.IP.1 DE.AE.1/DE.CM.7 ID.AM.6/ID.GV.2 PR.AT.2&5 **ISO 27002:2022:** 8.17/5.2 **ISF SOGP:** TI1.1/SM2.1	Evidence that the following have been completed: • Project Plans • Target Operating Models • Role Descriptions/Skills Matrices • Process Documentation and Technology
1.5	Ensure data and technical readiness prior to AI adoption, including high-quality, well-governed data, appropriate infrastructure, security controls and integration capabilities to support and manage AI technologies effectively.	**NIST CSF:** ID.AM.5 **ISO 27002:2022:** N/A **ISF SOGP:** N/A	Evidence of technical and data readiness for AI adoption
1.6	Establish a suitable AI governance framework or incorporate it into a suitable existing governance framework. This governance structure must clearly define AI governance principles, document acceptable AI usage policies, and oversee all AI risk management activities.	**NIST CSF:** ID.BE.3/ID.GV.4 ID.RM.3/PR.IP.7 **ISO 27002:2022:** N/A **ISF SOGP:** SG1.1	Governance ToRs Governance Meeting minutes, actions and decision logs

Reference	Minimum requirement	Control reference	Compliance Metric / Artefacts
1.7	Ensure policies and security standards are regularly reviewed and updated to reflect evolving AI technology and security threats.	**NIST CSF:** ID.AM.3/PR.DS.1/ PR.DS.5/ID.GV.1/ RS.CO.2 **ISO 27002:2022:** 5.01/5.12/6.04/8.12 **ISF SOGP:** IM1.5/SM1.1	Evidence of up-to-date cyber security policy and standards relating to AI
1.8	Clearly define roles and responsibilities for AI oversight, ensuring that accountability for AI-driven decisions is explicitly assigned to individuals or teams.	**NIST CSF:** ID.GV.2 **ISO 27002:2022:** N/A **ISF SOGP:** N/A	RACI Matrix Decision Review and Escalation processes Governance meetings minutes
1.9	AI technologies must support, not replace, human oversight, ensuring informed decision-making and accountability. • Keep human operators responsible for AI-assisted decisions, ensuring final accountability remains with personnel overseeing AI use. • Establish clear human checkpoints within AI-driven processes to review critical decisions, challenge AI outputs and apply human expertise where necessary.	**NIST CSF:** PR.AT.5 **ISO 27002:2022:** 5.2/5.3 **ISF SOGP:** SM2.1	Audit reports, decision-making processes and governance policies Risk assessments and ownership Evidence of defined roles and responsibilities

Reference	Minimum requirement	Control reference	Compliance Metric / Artefacts
1.10	Educate all staff and officers about the risks of AI technologies and the appropriate use of these tools, including where and how they are authorised to use them. Cybersecurity training must incorporate AI security content, which should be reviewed and updated regularly to address new threats.	**NIST CSF:** PR.AT.1&2/ PR.IP.11 **ISO 27002:2022:** 6.03/7.07 **ISF SOGP:** ST1.1/ST1.2/ST1.3	Education and Awareness (E&A) materials relating to the use of AI and E&A tracking and measurement of effectiveness of E&A campaign
1.11	Establish clear guidelines and documentation for prohibited AI use cases. Ensure end-users understand limitations and restrictions. • Use threat modelling to identify and inform users of all known harmful states, unmitigated risks, and misuse scenarios. • Implement controls to actively monitor, detect and prevent prohibited use cases.	**NIST CSF:** PR.AT.1 **ISO 27002:2022:** 5.1 **ISF SOGP:** SM1.2	Network monitoring logs End point management reports Low Level Designs
2	AI Data Security		
2.1	Ensure AI technologies comply with data protection policies. Consider factors such as data classification, risks associated with personal data and appropriate mitigations. This includes ensuring organisational data is not input into an AI-based or ML technologies without explicit consent.	**NIST CSF:** ID.GV.3/PR.DS.1&2 **ISO 27002:2022:** 5.1/5.31/5.34 8.11 **ISF SOGP:** IM1.4	Evidence that the following has been completed: • DPIA • DPO Engagement • Third Party Assurance Reports

Reference	Minimum requirement	Control reference	Compliance Metric / Artefacts
2.2	AI technologies must comply with data residency requirements, ensuring that all the data, including training datasets, operational data and outputs are stored and processed within approved jurisdictions.	**NIST CSF:** ID.SC.3/PR.IP.6 **ISO 27002:2022:** 5.14 **ISF SOGP:** IM1.5	Evidence that the following has been completed: • DPIA • Documented encryption and security controls • Low Level designs
2.3	Ensure AI technologies are limited to access only the data necessary for the specific tasks they are designed to perform. Unnecessary access increases security risks and exposes sensitive information. Access permissions should be regularly reviewed to ensure compliance with the principle of least privilege.	**NIST CSF:** PR.AC.1&4&6/ PR.PT.3 **ISO 27002:2022:** 5.15 **ISF SOGP:** AC1.1	Design decision logs Low Level Designs RBAC and Access Control policies
2.4	Access controls must be enforced to restrict who can read, input or amend data in the AI technologies to prevent unauthorised and unexpected changes.		
2.5	Verify that suppliers and open-source components meet existing security standards on identity management, access control and authentication. Ensure supply chain assurance extends to software and service providers involved in the development and maintenance of AI technologies.	**NIST CSF:** ID.SC.4 **ISO 27002:2022:** 5.19/5.21/5.22 **ISF SOGP:** SC1.4	Evidence that the following has been completed: • Third Party Assurance Reports • Legal/Regulatory guidance obtained. • Business Impact Assessments • Identified risks have been recorded in relevant risk registers.
2.6	Ensure third-party suppliers are assessed for any embedded use of AI within their services.		

Reference	Minimum requirement	Control reference	Compliance Metric / Artefacts
2.7	Mandate a risk assessment process for all AI components, covering governance, known risks, and the use of personal data. Require that all components be sourced from trusted and approved providers, with documented provenance details. Ensure that change management processes are in place that trigger risk assessment reviews. Triggers includes expansion or changes of use.	**NIST CSF:** ID.SC.1&2 ID.GV.3/PR.DS.5 **ISO 27002:2022:** 5.19/5.21/5.22 **ISF SOGP:** SC1.1/SC1.2	Evidence that the following has been completed: • Risk Assessments • Business Impact Assessments • Risk registers • DPIAs • Third Party Assurance Reports
2.8	Evaluate AI data sources to ensure they comply with security, privacy, and ethical requirements. Ensure suppliers provide transparency regarding data provenance and handling. Data should be sourced from trusted, verified sources, and any third-party data sets must meet the same standards for integrity and compliance.	**NIST CSF:** PR.IP.1&2 **ISO 27002:2022:** 5.12/5.13 5.14/5.33 **ISF SOGP:** IM1.2	Evidence that the following has been completed: • Risk Assessments • Risk registers • DPIA • Third Party Assurance Reports
2.9	Ensure that the intended usage of an AI technology is appropriate to the sensitivity of the data it was trained on as well as the controls intended to ensure the security of the data.		

Reference	Minimum requirement	Control reference	Compliance Metric / Artefacts
2.10	Ensure AI training data is securely protected to maintain the integrity and reliability of AI technologies: • Verify that training data is encrypted both at rest and in transit, with encryption protocols that meet industry standards. • Training data must be sourced from verified, ethical and legal sources. • Training datasets must account for the geographical, contextual, behavioural and functional elements relevant to the specific environment or setting in which the AI technology will be deployed. • Training data must be regularly tested to identify and mitigate bias. Require suppliers to provide proof of bias testing, along with evidence of mitigation measures and quality control processes. • Implement integrity checks and validation techniques to prevent poisoning attacks. • Ensure that pre-trained models (e.g., open source LLMs), commonly used as a base for AI technologies, are also assured for security risks.	**NIST CSF:** PR.DS.1&2 **ISO 27002:2022:** 5.14/5.33 5.35/8.16 **ISF SOGP:** IM1.2/SM1.5	Third Party Assurance Reports Evidence of supply chain analysis, Pen tests, ITHC reports Independent assessments/certifications (e.g., SOC 2 Type II, ISO/IEC 42001)

Reference	Minimum requirement	Control reference	Compliance Metric / Artefacts
3	AI Technical Security		
3.1	AI technologies must be kept up to date with regular security patches, updates, and performance improvements. Policing organisations must ensure that all components, including models, software, and underlying infrastructure, receive timely updates to protect against emerging threats and vulnerabilities.	**NIST CSF:** RS.AN.5/RS.MI.3 **ISO 27002:2022:** 8.8 **ISF SOGP:** TP2.2	Patch management processes Low level designs
3.2	Use vulnerability management tools and techniques to regularly monitor and address vulnerabilities within AI technologies. These tools should scan AI models, training data and deployment environments to identify at risk components and harden them against potential security breaches or compromise. Vulnerabilities must be patched promptly, and security controls should be strengthened to minimise risks.	**NIST CSF:** ID.RA.1&2/PR.IP.12 PR.IP.3/PR.DS.6 /DE.CM.8/RS.AN.5/ RS.MI.3 **ISO 27002:2022:** 8.8/8.18 **ISF SOGP:** TP2.1	Establish VM processes and policies VM tool reports VM remediation logs
3.3	Continuously update and maintain threat intelligence feeds to monitor for attacks or compromises of AI technologies.	**NIST CSF:** ID.RA.2&3/DE.AE.2 /RS.CO.5/RS.AN.5 **ISO 27002:2022:** 5.7 **ISF SOGP:** SE1.3	Evidence of Threat Intel feeds and monitoring and threat intel reports relating to AI

Reference	Minimum requirement	Control reference	Compliance Metric / Artefacts
3.4	AI generated content must be independently verified to confirm accuracy and prevent misinformation. To ensure outputs are reliable, policing organisations should: • Implement automated validation mechanisms. • Cross-check AI outputs against reliable sources. • Establish human oversight processes to ensure that AI decisions are accurate and fair.	**NIST CSF:** N/A **ISO 27002:2022:** 5.35/8.16 **ISF SOGP:** AS1.1	Evidence of Test reports, Audit Logs, Low level design documents Third Party Assurance reports
3.5	AI technologies must provide clear, understandable and consistent explanations for decisions made. Therefore, policing organisations should: • Evaluate AI technologies against diverse user groups to identify potential biases and disparities in decision-making. • AI technologies must allow users to challenge, review and override AI outputs. • Establish internal procedures to regularly review and audit decision making to ensure output remains accurate.	**NIST CSF:** DE.DP.5 **ISO 27002:2022:** 5.35 **ISF SOGP:** AS2.1/AS2.2	Evidence of Test reports Audit Logs Low level designs Third Party Assurance reports

Reference	Minimum requirement	Control reference	Compliance Metric / Artefacts
3.6	AI technologies must be protected against adversarial attacks that can manipulate their behaviour, leading to incorrect or malicious outputs. To mitigate these risks: • Regularly perform penetration tests to identify and address vulnerabilities in the AI technologies. For third-party systems, obtain and review recent test results. • Validate and sanitise all incoming data to prevent malicious input from compromising the model. • Continuously monitor outputs for signs of unexpected or manipulative behaviour, with alerts for suspicious activity. • Conduct resilience testing and implement mitigation strategies accordingly.	**NIST CSF:** DE.AE.2&3&4&5 DE.CM.6/DE.DP.5 **ISO27002:2022:** 5.25/8.15/8.16 **ISF SOGP:** SE1.2	Low Level Designs Evidence of supply chain analysis, ITHC/Pen Test reports

Reference	Minimum requirement	Control reference	Compliance Metric / Artefacts
3.7	AI data sources, including external supplier feeds, live operational data and datasets, can be actively targeted to undermine model reliability. To mitigate these risks, organisations should: • Regularly audit and validate AI data sources to spot anomalies, inconsistencies or suspicious patterns. • Establish alerts for unexpected data types or results. • Ensure clear records of data origin and any modifications to track and verify AI data integrity. • Scrutinise the supply chain for gaps that attackers can insert themselves into. • Regularly assure connections with external suppliers for any opportunities for model loss, damage, or pollution of the dataset.	**NIST CSF:** PR.AT.3/DE.AE.1&2 DE.CM.7/RS.AN.1 **ISO27002:2022:** 5.14/5.2/8.16 **ISF SOGP:** SR1.1/TP1.3	Evidence of data sampling Low Level Designs Evidence of supply chain analysis, ITHC/Pen Test reports

Reference	Minimum requirement	Control reference	Compliance Metric / Artefacts
AI Business Continuity			
3.8	Clarify business impact of AI disruption (e.g. outages) and design for the appropriate levels of uptime and resiliency. Ensure the robustness of the platforms supporting the AI technology meet the levels of uptime and resiliency required.	**NIST CSF:** PR.AC.4/ID.BE.5/ PR.PT.5 **ISO 27002:2022:** 5.3/8.27 **ISF SOGP:** BC2.1	Evidence that the following has been completed: • Business Impact Assessment • Risk Register • Service Level Agreements • Low Level Designs • Design Decision Logs
3.9	Update disaster recovery plans to address AI-specific risks, including adversarial attacks and data poisoning, ensuring readiness for prompt recovery. Ensure Business Continuity Plans consider during service disruption, the ability of personnel to perform tasks undertaken by AI.	**NIST CSF:** PR.IP.9&10 **ISO 27002:2022:** 5.3 **ISF SOGP:** BC2.2/BC2.3	Business continuity plans ensure that personnel can perform tasks that were automated by AI. Evidence of updated and tested Disaster Recovery and Business Continuity plans

Reference	Minimum requirement	Control reference	Compliance Metric / Artefacts
3.10	To minimise service disruption and reduce AI technology exposure to threats, policing organisations should: • Implement appropriate levels of network segregation and access control to prevent lateral movement and direct access. • Deploy AI models in isolated environments. Only use or deploy a dedicated tenant-specific model, avoiding the use of shared models that may be trained on data from multiple tenants. • Use threat intelligence and modelling techniques to assess current threat levels and anticipate emerging risks. • Use red team testing techniques to explore possible attack patterns. • Limit AI model exposure to internal systems and manage how much information the AI model returns to queries to limit the ability of threat actors to gather useful attack data.	**NIST CSF:** PR.IP.1/ PR.AC.1&4&5&6 /PR.PT.3&4/ ID.RA.2&3/DE.AE.2 /RS.CO.5/RS.AN.5/ DE.AE.3/ DE.CM.1&3&3&7/ RS.RP.1/RS.IM.1 **ISO 27002:2022:** 8.09/8.20/5.07 /5.15 **ISF SOGP:** NC1.1/SE1.3/AC1.1	Evidence that the following has been completed: • Business Impact Assessment • Low Level Designs • Design Decision Logs • Threat Intelligence Reports • Red Team Reports

Reference	Minimum requirement	Control reference	Compliance Metric / Artefacts
4	AI Incident Response and Monitoring		
4.1	All AI decisions and interactions must be securely logged with tamper resistant storage to prevent unauthorised modifications. Logs must be encrypted, stored immutably, and reviewed regularly for anomalies. Suppliers must confirm that their logs are comprehensive, include critical interactions and meet encryption and tamper-resistance standards. Access to logs must be carefully controlled through role-based access, ensuring that only authorised personnel can view log data, based on their role and the level of data sensitivity.	**NIST CSF:** PR.DS.5 **ISO 27002:2022:** 5.16/5.18 8.15 **ISF SOGP:** SE1.1	Design decision log Low Level Design documents Security Operations Centre Playbooks
4.2	Establish a dedicated AI-specific incident response plan to address the unique challenges of AI-related attacks. This plan should include clearly defined workflows for incident management, including rollback and shutdown procedures, as well as rapid deactivation protocols for swift risk mitigation. To ensure effectiveness, incident response strategies should be regularly tested through tabletop exercises and AI-specific playbooks should be developed to guide incident handling.	**NIST CSF:** PR.IP.2/PR.IP.10 DE.AE.2&3&4&5 /DE.CM.1&6&7 /DE.DP.2&4&5 /RS.AN.1 **ISO 27002:2022:** 5.25/8.15/8.16 **ISF SOGP:** SE1.2/SR1.6	Evidence of an AI-specific Incident Response Plan (IRP) IRP Test Plans IRP Test results IRP improvement plans Security Operations Centre Playbooks

Reference	Minimum requirement	Control reference	Compliance Metric / Artefacts
4.3	All AI security incidents must be logged, reported, and analysed. Suppliers must confirm that detailed incident analysis reports are generated and shared with affected customers.	**NIST CSF:** PR.IP.9/RC.IM.1 RC.IM.2/RC.CO.3 **ISO 27002:2022** 5.25/8.15/8.16 **ISF SOGP:** SE1.2	Incident response plans reports, post incident reviews and improvement plans
4.4	Develop tailored logging and monitoring requirements for each AI instance, ensuring continuous tracking of AI behaviour and data inputs/outputs. This includes monitoring for anomalous behaviour, signs of manipulation and potential security breaches. Logs should be reviewed regularly for suspicious activities and real time alerts should be set up for immediate responses to any incidents.	**NIST CSF:** PR.PT.1/DE.AE.1&3/ DE.CM.1&3&7/ DE.DP.2&4/RS.CO.2 /RS.AN.1 **ISO 27002:2022:** 8.15/8.17 **ISF SOGP:** SE1.1	Design decision logs Low Level Designs Security Operations Centre Playbooks AI monitoring use cases
4.5	All AI-related assets must be recorded in the relevant asset register. Each record should clearly specify the asset's approved use, ownership, and any other information relevant to the AI technology and its usage. The asset register should be regularly reviewed and updated to ensure accuracy and completeness.	**NIST CSF:** ID.AM.1&2 **ISO 27002:2022:** 5.2/5.9 **ISF SOGP:** AM1.1	Technology Asset Inventory

Communication approach

This standard will be communicated as follows:

1. Internal peer review by the members of the National Cyber Policy and Standard Working Group (NCPSWG), which includes representatives from PDS and participating forces.
2. Presentation to the National Cyber Policy & Standards Board (NCPSB) for approval.
3. Formal publication and external distribution to PDS community, police forces and associated bodies.

This standard should be distributed within IT and project teams to help complete an initial gap analysis which can inform any implementation plan. This implementation plan can be shared with force SIROs / Security Management Forum / Information Management. Consideration should also be given to raising awareness amongst force personnel of the implementation of this standard where it may affect them.

This standard should be mapped to a project lifecycle and internal governance prior to adoption. Following this, it should be provided to the Information Assurance communities and PMO's and should also be shared with procurement & commercial leads to ensure this is built into procurement activities.

Measurables generated by adopting this standard can also form part of regular Cyber management reporting and audit evidencing.

Review Cycle

This standard will be reviewed at least annually (from the date of publication) and following any major change to Information Assurance (IA) strategy, membership of the community, or an identified major change to the cyber threat landscape. This ensures IA requirements are reviewed and that the standard continues to meet the objectives and strategies of the police service.

Document Compliance Requirements

(Adapt according to Force or PDS Policy needs.)

Equality Impact Assessment

(Adapt according to Force or PDS Policy needs.)

Document Information

Document Location

https://knowledgehub.group/web/national-standards/policing-standards

Revision History

Version	Author	Description	Date
0.1	PDS Cyber	Initial draft	06/03/25
0.2	PDS Cyber	Updated to incorporate internal peer review comments	07/05/25
1.0	PDS Cyber	Updated to incorporate NCPSWG review comments	07/05/25

Approvals

Version	Name	Role	Date
1.0	NCPSB	National Cyber Policy & Standards Board	22/05/25

Document References

Document Name	Version	Date
ISF - Standard of Good Practice (for Information Security)	v2024	03/2025
ISO 27002:2022 - Information security, Cybersecurity and privacy protection – Information security controls	v2022	03/2025
CIS Controls	v8	05/2021
NIST Cyber Security Framework	v1.1	04/2018
CSA Cloud Controls Matrix	v4	01/2021
10 Steps to Cyber Security - NCSC.GOV.UK	Web Page	05/2021
ai_principles_1_1_1.pdf	v1.1.1	03/2025
AI Playbook for the UK Government - GOV.UK	-	03/2025
Code of Practice for the Cyber Security of AI - GOV.UK	-	03/2025
Artificial Intelligence Toolkit (interpol.int)	-	03/2025
Principles for security of Machine learning ML - NCSC.GOV.UK	v2.0	03/2025
Guidelines for secure AI system development - NCSC.GOV.UK	v1.0	03/2025
OWASP Top 10 for Large Language Model Applications | OWASP Foundation	v2025	03/2025