TrustIoT Framework for Industry 4.0

"AI Training & Data Protection"

|  |  |
| --- | --- |
| Document Classification: | Internal |
| Document Ref. | *TrustIoT Framework for Industry 4.0* |
| Version: | *1* |
| Document Author: | *Jibran Saleem* |
| Document Owner: |  |

**Revision History**

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Revision Author** | **Summary of Changes** |
|  |  |  |  |
|  |  |  |  |

**Distribution**

|  |  |
| --- | --- |
| **Name** | **Title** |
|  |  |
|  |  |
|  |  |

**Approval**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Position** | **Signature** | **Date** |
|  |  |  |  |

Table of Contents

[1. Introduction 4](#_Toc190905581)

[2. Purpose 4](#_Toc190905582)

[3. Scope 4](#_Toc190905583)

[4. Policy Statement 4](#_Toc190905584)

[4.1. Data Collection and Storage 4](#_Toc190905585)

[4.2. Data Access and Control 4](#_Toc190905586)

[4.3. Data Anonymisation and Pseudonymisation 4](#_Toc190905587)

[4.4. Data Integrity and Validation 5](#_Toc190905588)

[4.5. Bias and Fairness 5](#_Toc190905589)

[5. Responsibilities 5](#_Toc190905590)

[6. Breaches of Policy 5](#_Toc190905591)

[7. Document Management 5](#_Toc190905592)

# Introduction

Artificial Intelligence (AI) and Machine Learning (ML) models are increasingly integrated into IoT systems to enable advanced analytics, automation, and decision-making capabilities. These models rely heavily on data for training and refinement. Ensuring the security, privacy, and integrity of this training data is paramount to protect sensitive information, maintain the trustworthiness of AI outputs, and prevent potential biases. This policy outlines the measures to be taken to safeguard IoT data used for AI training purposes.

# Purpose

The purpose of this policy is to establish a framework for protecting the confidentiality, integrity, and availability of IoT data used for AI and ML model training within the organisation. This policy aims to:

* Ensure that IoT training data is classified, handled, and protected in accordance with its sensitivity and applicable regulations.
* Restrict access to IoT training data to authorised personnel and systems only.
* Implement appropriate security measures to protect data at rest and in transit.
* Minimise the collection and retention of unnecessary data.
* Mitigate the risk of data poisoning and bias in AI models.

# Scope

This policy applies to all data generated, collected, processed, or stored by IoT devices and systems within the organisation that is used for AI and ML model training purposes.

# Policy Statement

## Data Collection and Storage

* **Data Collection:** IoT training data shall be collected in a lawful and transparent manner, with clear notice and consent mechanisms in place where required by applicable regulations.
* **Data Classification:** Training data shall be classified based on its sensitivity and the potential impact of unauthorised access or disclosure. Appropriate data protection measures shall be applied based on the data classification.
* **Secure Storage:** Training data shall be stored in secure environments, utilising encryption and access controls to protect against unauthorised access, modification, or exfiltration.
* **Data Retention:** Training data shall be retained only for as long as necessary for the purposes of model development and training, or as required by law.

## Data Access and Control

* **Role-Based Access Control (RBAC):** Access to IoT training data shall be granted based on user roles and responsibilities within the organisation, utilising RBAC or similar mechanisms.
* **Need-to-Know Basis:** Access shall be granted on a need-to-know basis, limiting exposure to sensitive data.
* **Data Sharing Agreements:** Sharing of IoT training data with external parties shall be governed by formal agreements that outline data protection and confidentiality requirements.

## Data Anonymisation and Pseudonymisation

* **Privacy Enhancement:** Where feasible, data anonymisation or pseudonymisation techniques shall be employed to minimise the risk of identifying individuals from training data.
* **Re-identification Risk:** The risk of re-identification shall be assessed before using anonymised or pseudonymised data for training, and appropriate safeguards shall be implemented to mitigate this risk.

## Data Integrity and Validation

* **Data Integrity Checks:** Mechanisms shall be implemented to verify the integrity of training data and detect any unauthorised modifications or tampering.
* **Data Validation:** Training data shall be validated to ensure its accuracy, completeness, and consistency before being used for model development.
* **Data Cleansing:** Data cleansing techniques shall be applied to remove errors, inconsistencies, or irrelevant data that may impact model performance or introduce bias.

## Bias and Fairness

* **Bias Assessment:** Training data and AI models shall be assessed for potential biases that may lead to discriminatory or unfair outcomes.
* **Bias Mitigation:** Measures shall be taken to mitigate or address identified biases in training data or model algorithms.
* **Fairness:** AI models shall be designed and trained to promote fairness and avoid discriminatory outcomes.

# Responsibilities

* **Information Security Officer:** Responsible for overseeing the implementation and enforcement of this policy.
* **Data Scientists and AI Developers:** Responsible for handling IoT training data in accordance with this policy and implementing appropriate data protection measures.
* **IT Department:** Responsible for providing secure infrastructure and tools for data storage, access control, and anonymisation/pseudonymisation.
* **Data Owners:** Responsible for classifying data and defining access control policies.

# Breaches of Policy

Non-compliance with this policy may result in disciplinary action, up to and including termination of employment or contractual relationships. Additionally, breaches of data protection regulations may result in legal and financial penalties for the organiation.

# Document Management

This document is valid as of [dd/mm/yyyy].

This document is reviewed periodically and at least annually to ensure compliance with the following prescribed criteria.

* Compliant with the Internet of Things (IoT) Security Framework for Industry 4.0.
* Legislative requirements defined by law, where appropriate.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

[Name 1]

Manager