HUMBERSIDE FIRE AND RESCUE SERVICE

Service Improvement

Artificial Intelligence (AI) Policy

Owner	**Executive Director of Corporate Services**
Responsible Persons	**Head of Corporate Assurance** **Head of Corporate Risk & Intelligence** **Head of Digital Services**
Date Written	**December 2023**
Date of Last Review	**July 2024**
Date of next review	**July 2025**
EIA Completed	**December 2023**

What we must do well

How we support our communities

We value and support the people we employ

We efficiently manage the Service

CONTENTS

1. INTRODUCTION

Artificial Intelligence (AI) can be defined as the use of digital technology to create systems capable of performing tasks commonly thought to require human intelligence.

At its core, AI is a research field spanning philosophy, logic, statistics, computer science, mathematics, neuroscience, linguistics, cognitive psychology, and economics.

AI is constantly evolving but generally involves machines using statistics to find patterns in large amounts of data. AI can perform repetitive tasks with data without the need for constant human guidance.

The purpose of this Policy is to provide a framework for the use of AI applications by the Service. The Policy is designed to ensure that the use of AI is ethical, complies with all applicable laws, regulations, and Service policies, and complements the Service's existing Information Security Policy.

The pace of development and application of AI is such that this Policy will be in a constant state of development.

Core Code of Ethics

HFRS has adopted the Core Code of Ethics for Fire and Rescue Services. The Service is committed to the ethical principles of the Code and strives to apply them in all we do, therefore, those principles are reflected in this Policy.

National Guidance

Any National Guidance which has been adopted by HFRS will be reflected in this Policy.

2. EQUALITY, DIVERSITY & INCLUSION

HFRS has a legal responsibility under the Equality Act 2010, and a commitment, to ensure it does not discriminate either directly or indirectly in any of its functions and services or in its treatment of staff, in relation to race, sex, disability, sexual orientation, age, pregnancy and maternity, religion and belief, gender reassignment or marriage and civil partnership. It also has a duty to make reasonable adjustments for disabled applicants, employees and service users.

3. AIM AND OBJECTIVES

- To ensure that the Service has the necessary safeguards in place to protect the rights and freedoms of individuals when AI is used.

- That both the Service and its employees operate in compliance with data protection principles and law and other relevant legislation and cyber security measures when employing AI.

- That employees are aware of their responsibilities and associated risks when using AI applications.

4. ASSOCIATED DOCUMENTS

- Equality Impact Assessment

- Legal References
  - Data Protection Act 2018
  - UK General Data Protection Regulation
  - *A Guide to Using Artificial Intelligence in the Public Sector* (Office for Artificial Intelligence & Government Digital Service - January 2020)
  - *National AI Strategy* (HM Government) – September 2021
  - *Guidance on AI and Data Protection* (Information Commissioner's Office) – March 2023
  - Cyber Security Policy
  - Data Protection Policy
  - Privacy and Consent Policy Delivery Guidance
  - Data Protection Impact Assessment Policy Delivery Guidance
  - Data Protection Request Policy Delivery Guidance (inc. SAR)
  - Personal Data Breach Notification Policy Delivery Guidance
  - Information Security Policy
  - Internet, E-Mail and Instant Messaging Policy
  - Records Management and Data Quality Policy
  - Disposal and Destruction of Records Policy Delivery Guidance
  - AI Request Form

The following sources have been used to help inform this Policy:

- *A guide to using AI in the public sector* (Office for Artificial Intelligence & Government Digital Service) – January 2020
- *A Pro-innovation approach to AI Regulation* (Department for Science, Innovation & Technology) – March 2023
- *Guidance on AI and Data Protection* (Information Commissioner's Office) – March 2023
- *Generative AI: eight questions that developers and users need to ask* (Information Commissioner's Office) – April 2023
- *GenAI Corporate Policy Sample Framework* (ALGIM & Socitm) – June 2023
- *How to implement Good Information Governance into Artificial Intelligence & Machine Learning Projects* (Act Now Training) – July 2023

- National Guidance
  There is no specific national guidance relating to this policy

5. DEFINITIONS

For this Policy, the following definitions relate:

- **Artificial Intelligence (AI)**
  The theory and development of computer systems able to perform tasks normally requiring human intelligence, such as text and image generation, visual perception (including facial recognition), speech recognition, decision-making and translation between languages.

- **Machine learning**
  A subset of AI that refers to the use and development of computer systems that are able to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyse and draw inferences from patterns in data. There are two main types of machine learning: supervised learning and unsupervised learning.

    - **supervised learning** allows an AI model to learn from 'labelled' or 'known' data.
    - **unsupervised learning** is the training of an AI algorithm to use unlabelled and unclassified information to learn for itself.
        - **reinforcement (deep) learning** allows an AI model to learn as it performs a task (like a human would) and change its course.

- **Generative artificial intelligence (GenAI)**
  Generative artificial intelligence (also generative AI or GenAI) 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. Generative AI starts with a prompt that could be in the form of a text, an image, a video, a design, musical notes, or any input that the AI system can process. Various AI algorithms then return new content in response to the prompt.

6. USE OF GenAI TECHNOLOGY (SUCH AS CHATBOTS AND VIDEO/PHOTO CREATION)

External GenAI Technology

This policy applies to all users who wish to access external GenAI, whether through Service-owned devices or BYOD (bring your own device) in pursuit of Service activities and must not contravene the Information Security Policy and Internet, Email and Instant Messaging Policy.

GenAI must be used in a manner that promotes fairness and avoids bias to prevent discrimination and promote equal treatment and be used in such a way as to contribute positively to the Service's goals and values.

Due to the inherent risks associated in using external GenAI, employees must employ extreme caution when using such technology. However, it is recognised

that certain GenAI can be a useful professional tool (for example by generating text or content for reports, emails, presentation or Service communications) if used appropriately. An employee must assure themselves that **all** of the following directives can be met, then it is permissible to use such technology. **No personal data should be entered into an external GenAI application.**

- **Risk and Security**
  Use of GenAI carries inherent risks. GenAI may store sensitive data and information, which could be at risk of being breached or hacked. The Service must assess the technical protections and security certification of GenAI before use. The threat to cyber security through the use of AI applications, such as chatbots, is real. Cybercriminals may be redirecting their focus to crafting more sophisticated scams through AI chatbots that exploit user trust (for example through phishing scams). **If a user has any doubt about the security of information input into GenAI, they should not use GenAI.**

- **Copyright**
  Users must adhere to copyright laws when utilising GenAI. It is prohibited to use GenAI to generate content that infringes upon the intellectual property rights of others, including but not limited to, copyrighted material. **If a user is unsure whether a particular use of GenAI constitutes copyright infringement, then it should not be used.**

- **Accuracy**
  All information generated by GenAI must be reviewed and edited for accuracy prior to use. Users of GenAI are responsible for reviewing output and are accountable for ensuring the accuracy of GenAI-generated output before use/release. **If a user has any doubt about the accuracy of information generated by GenAI, they should not use GenAI.**

- **Confidentiality**
  Confidential, commercially sensitive, or personal information must not be entered into any GenAI tool, as this information may enter the public domain. Users must follow all applicable data privacy laws and Service policies when using GenAI. If pseudonymised data is input, the user must be certain that AI output does not then allow reidentification of individuals. **If a user has any doubt about the confidentiality of information, they should not use GenAI.**

- **Ethical Use**
  GenAI must be used ethically and in compliance with all applicable legislation, regulations, and Service policies. Users must not use GenAI to generate content that is discriminatory, offensive, or inappropriate. **If there are any doubts about the appropriateness of using GenAI in a particular situation, users should consult with their Information Asset Owner.**

- **Disclosure**
  Content produced via GenAI must be identified and disclosed as containing GenAI-generated information.

Footnote example: ***Note:*** *This document contains content generated by Artificial Intelligence (AI). AI-generated content has been reviewed by the author for accuracy and edited/revised where necessary. The author takes responsibility for this content.*

- **Legal Compliance**
  Data entered into GenAI may enter the public domain. This can release non-public information and breach regulatory requirements, customer or vendor contracts, or compromise intellectual property. Any release of private/personal information without the authorisation of the information's owner could result in a breach of relevant data protection laws. The use of GenAI to compile content may also infringe on regulations for the protection of intellectual property rights. **Users should ensure that their use of any GenAI complies with all applicable laws and regulations and with Service policies.**

- **Bias and Discrimination**
  GenAI may make use of and generate biased, discriminatory or offensive content. **Users should use GenAI responsibly and ethically, in compliance with Service policies and applicable laws and regulations.**

- **Data Sovereignty and Protection**
  While a GenAI platform may be hosted internationally, under data sovereignty rules information created or collected in the originating country will remain under the jurisdiction of that country's laws. The reverse also applies. If information is sourced from GenAI hosted overseas, the laws of the source country regarding its use and access may apply. **GenAI service providers should be assessed for data sovereignty practice by any individual wishing to use their GenAI.**

Internal GenAI technology

The use/development of internal chatbots should limit the amount of personal data used. The dataset used to feed the chatbot and its algorithm should only be sourced from Service owned data to ensure the accuracy and validity of the data. However, if the only viable option is to use an external dataset, this must be sourced from reputable and trusted sources whose data can be validated.

A comprehensive risk assessment should be conducted, against all of the above areas and those detailed in Section 7, for any project or process where use of GenAI is proposed, using the Risk Matrix and Impact Assessment as part of the Strategic Risk and Opportunity Register.

7. USE OF AI MACHINE LEARNING CONSIDERATIONS

With any AI project you must consider several factors. A comprehensive risk assessment must be conducted, against all of the following points and any relevant aspects detailed in Section 6, using the Risk Matrix and Impact Assessment as part of the Strategic Risk and Opportunity Register:

- **Data Provenance -** It is essential to know the origin and history of the data used in the context of AI because it can affect the accuracy and reliability of the resulting AI system. If data is not collected or labelled properly, the AI system may underperform or generate errors or biased results. **If there are any doubts about the provenance of the data, it should not be used. Actions can be taken to improve the data or identify better data for the purpose.**

- **Fairness** - are the models trained and tested on relevant, accurate, and generally applicable datasets and is the AI system deployed by users trained to implement them responsibly and without bias.

- **Accountability** - consider who is responsible for each element of the model's output and how the designers and implementers of AI systems will be held accountable. If the machine learning AI requires supervision, then a change management programme is required to be put in place to continually review and maintain the algorithm used so that any 'model drift' can be accounted for and version controlled.

  If at any point 'model drift' becomes apparent in the AI algorithm, there must be a clear audit trail as to how this has changed (see 'Explainability and transparency'). The new AI algorithm must be reviewed to ensure it is still in scope of the initial concept and if necessary, the relevant privacy notice reviewed and amended accordingly. Such changes in AI Algorithms must be reported and justified by the programmer/relevant employee to the relevant Information Asset Owner.

  If a data subject submits a subject access request (SAR) then the Service must be confident it can track all the stages through which that person's data is being processed. It must still be able to respond to people's requests for access, rectification, erasure or other information rights.

- **Mitigate security risks** - in addition to personal data leakage risks, you should consider and mitigate risks of model inversion and membership inference, data poisoning and other forms of adversarial attacks.

- **Limit unnecessary processing** - you must collect only the data that is adequate to fulfil your stated purpose. The data should be relevant and limited to what is necessary.

- **Privacy** - complying with appropriate data policies, for example, the General Data Protection Regulation (UK GDPR) and the Data Protection Act 2018.

- **Explainability and transparency** - Explainability is the idea that a machine learning model and its output can be explained in a way that is comprehensible to a human being at an acceptable level. In other words, the affected stakeholders must be able to know how the AI model reached its decision.

- **Costs** - consider how much it will cost to build, run and maintain an AI infrastructure, train and educate staff and if the work to install AI may outweigh any potential savings/improved intelligence.

Always be mindful that if personal data, as defined under GDPR, is being processed within an algorithm/AI model/application then data protection legislation and principles apply.

There are several 'off the shelf' data sets available for purchase to build algorithms on. Only trusted and reputable datasets should be used. It is the Information Asset Owner's (IAO) responsibility to ensure the nature of the source of the dataset is genuine and ethically sourced.

Privacy Notices

The relevant Privacy Notice must be reviewed, amended appropriately and published prior to any AI project commencing.

Records Management, Data Protection Impact Assessment (DPIA) and Equality Impact Analysis (EIA)

Any AI application/tool developed and owned by an individual/Function must be referenced within the relevant Function's Record of Processing Activities (ROPA). A data protection impact assessment (DPIA) and Equality Impact Analysis (EIA) must be undertaken when proposing/developing the use of any type of AI.

Information Sharing Agreements and Contracts

Any contract or sharing agreement that utilises AI must ensure relevant clauses are inserted that testify to the information and data use, storage and security.

8. PROHIBITED USE OF AI

The Service prohibits employees or Functions to use Artificial Intelligence to undertake the following activities:

- **Automated decision making** (decisions made without human intervention, which have legal or similarly significant effects on data subjects).

- **Image recognition/computer vision** (the ability of a machine or programme to emulate human vision).

9. REQUESTS FOR APPROVAL OF AI USE

The development of an internal AI application/Machine Learning tool/model must not be commenced without the prior approval of the AI Working Group.

Requests **must** be made through the submission of a 'Use of AI Request Form' (Microsoft Form).

If the project is of a **strategic** nature, this **must** also be progressed under the Strategic Project Register process via the Service Improvement Supervisor in Corporate Assurance.

10. RECORD OF AI PROJECTS AND APPLICATIONS

The Service (through Corporate Assurance) will maintain a record of all AI tools/ applications proposed by individual employees and a record of each Function's owned AI application/tool.

Employees/Functions must keep Corporate Assurance informed of any changes to, or additional uses of, AI.

11. COMPLIANCE

Failure to comply with this policy may result in disciplinary action, in accordance with the Service's Disciplinary Procedure Policy.

If you require any further information regarding this policy, please contact Corporate Assurance.