Committee(s): Digital Services Committee	**Dated:** 6th November 2023
Subject: Generative AI Standard Operating Procedure	**Public**
Which outcomes in the City Corporation's Corporate Plan does this proposal aim to impact directly?	1
Does this proposal require extra revenue and/or capital spending?	**No**
If so, how much?	
What is the source of Funding?	
Has this Funding Source been agreed with the Chamberlain's Department?	
Report of: Gary Brailsford-Hart, Director of Information	**For Information**
Report author: Gary Brailsford-Hart, Information Management Services Director	

Summary

Generative Artificial Intelligence (GenAI) is an emerging technology that exposes the organization to both opportunity and risk. In response, we have drafted a standard operating procedure (SOP) to assist our staff in how to safely interact and use this technology all the while ensuring that this is undertaken within an ethical framework of principles.

The attached SOP is the first in a series of procedures and policies that will form the Data Ethics Framework for the organization.

Recommendation

Members are asked to:

- Note the attached procedure.

Main Report

Background

1. The rapid growth of Artificial Intelligence (AI) is unsurprising. The speed and accuracy that AI can bring to corporate processes make it an attractive way to deliver an effective and efficient service. However, the application of AI can be contentious[1]. Transparency and fairness must be at the heart of what we implement, to ensure a proportionate and responsible use that builds public confidence.

2. The Use of Generative Artificial Intelligence procedure defines how we should use this technology as well as outlining a set of principles that define how we should use AI in our business.

3. This procedure forms part of a wider framework being developed to describe and manage the ethical use of AI across the City.

Current Position

4. The building interest in the use and exploitation of AI is understood and without a clear set of policies and procedures we risk the organization utilizing these new technologies without the necessary oversight and control. Therefore, this procedure has been written to address the immediate needs and will form part of a more significant framework approach as we mature in this area.

Conclusion

5. This report has been produced to provide oversight on the direction of travel and control mechanisms being put in place to manage this emergent technology. Whilst we recognize the risks, we also recognise the significant opportunities this can deliver. We are on a journey with this technology and this paper represents one of our first steps in recognizing and defining how we will use it to deliver positive outcomes for the City.

Appendices

- Appendix 1 – Generative Artificial Intelligence Standard Operating Procedure

Gary Brailsford-Hart

Director of Information, City of London Police

T: 0207 601 2352
E: gary.brailsford@cityoflondon.police.uk

Appendix 1 – Generative Artificial Intelligence Standard Operating Procedure

Generative Artificial Intelligence

Standard Operating Procedure

2023

Version	Date	Comments
0.1	July 2023	Initial draft by Sam Collins
0.2	November 2023	Changed to SOP in order to be included within a broader AI & Ethics Policy and added the principles of AI as an appendix.
0.3		
0.4		

1. Purpose

The purpose of this Standard Operating Procedure (SOP) document is to provide a framework for the use of Generative Artificial Intelligence Large Language Models (GenAI) such as ChatGPT, Bard, Bing or other similar tools by City of London Corporation (COL) employees, contractors, developers, vendors, temporary staff, consultants or other third parties, hereinafter referred to as 'users'.

This policy is designed to ensure that the use of GenAI is ethical, complies with all applicable laws, regulations and policies, and complements existing COL information and security policies.

The pace of development and application of GenAI is such that this policy will be subject to regular review.

2. Use

This SOP applies to all users with access to GenAI, whether through COL-owned devices or BYOD (bring your own device) in pursuit of COL activities.

Use of GenAI must be in a manner that promotes fairness and avoids bias to prevent discrimination and promote equal treatment, and be in such a way as to contribute positively to COL goals and values.

Users may use GenAI for work-related purposes subject to adherence to the following policy. This includes tasks such as generating text or content for reports, emails, presentations, images and customer service communications.

Particular attention should be given to Governance, Vendor practices, Copyright, Accuracy, Confidentiality, Disclosure and Integration with other tools.

3. Governance

Before accessing GenAI technology, users must first notify the City of London Corporation's Information Management Board of their intention to use, the reason for use, and the expected information to be input as well as the generated output and distribution of content.

4. Vendors

Any use of GenAI technology in pursuit of COL activities should be done with full acknowledgement of the policies, practices, terms and conditions of developers/vendors.

5. 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, they should contact the Comptroller and City Solicitor's Department before using GenAI.**

6. 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.**

7. Confidentiality

Confidential information and personal data must not be entered into an GENAI tool, as information may enter the public domain. Users must follow all Data Protection principles as outlined in the Data Protection Act 2018 and the UK General Data Protection Regulation 2021 and organisational policies when using GenAI. **If a user has any doubt about the confidentiality of information, they should not use GenAI.**

8. Ethical Use

GenAI must be used ethically and in compliance with all applicable legislation, regulations and organisational 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 the Data Protection Team.**

9. 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.*

10. Integration with other tools

API and plugin tools enable access to GenAI and extended functionality for other services to improve automation and productivity outputs. API and plugin tools must be rigorously tested and approved prior to use, for:

- Moderation – to ensure the model properly handles hate, discriminatory, threatening, etc. inputs appropriately.
- Factual responses – provide a ground of truth for the API and review responses accordingly.

11. Risks

Use of GenAI carry inherent risks. A comprehensive risk assessment should be conducted for any project or process where use of GenAI is proposed. The risk assessment should consider potential impacts including: legal compliance; bias and discrimination; security (including technical protections and security certifications); and

data sovereignty and protection. Where the risk is of processing personal data then a Data Protection Impact Assessment will also need to be completed.

12. Legal compliance

Data entered into GenAI should be considered as being made public. 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. 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 COL policies.**

13. 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 COL policies and applicable laws and regulations.**

14. Security

GenAI may store sensitive data and information, which could be at risk of being breached or hacked. COL will ensure a technical assessment is conducted as part of the approval process to assess technical protections and security certification of GenAI before use. **If a user has any doubt about the security of information input into GenAI, they should not use GenAI.**

15. Data sovereignty and protection

While an GenAI platform may be hosted internationally, information created or collected in the United Kingdom of Great Britain and Northern Ireland (UK), under data sovereignty rules, is still under jurisdiction of UK laws. The reverse also applies. If information is sourced from GenAI hosted overseas for use in the UK, 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 organisation wishing to use their GenAI.**

16. Compliance

Any violations of this policy should be reported to IT, or where personal data is involved, the Data Protection Team. In every case senior management should be informed. Failure to comply with this SOP may result in disciplinary action, in accordance with COL HR policies and procedures.

17. Review

This policy will be reviewed periodically and updated as necessary to ensure continued compliance with all applicable legislation, regulations and organisational policies.

18. Acknowledgment

By using GenAI, users acknowledge that they have read and understood these guidelines, including the risks associated with the use of GenAI.

ANNEX A – Principles of Artificial Intelligence

Our AI principles are founded on three sets of guidance: the FAST Principles[2], the OECD AI Principles[3], and the Data Ethics Framework[4]. We apply these to CoL with the intent to support an openness to scrutiny, integrity, and public confidence in our use of AI technologies.

Principle A. Lawful: All use of AI will comply with applicable laws, standards, and regulations. This includes all users of AI, ML, ADA and related data processing ensuring the use is recorded centrally in the Data Protection Record of Processing Activity (ROPA).

Principle B. Transparent: All use of AI will be subject to 'Maximum Transparency by Default' (MTbD).

B1. We should ensure the public are aware of AI uses. This will typically include publishing an overview of the algorithms used and the known limitations of the training data used. The datasets will be present on the force IAR with allocated Information asset owners.

B2. Where operational or security requirements restrict the ability to share, the AI will undergo scrutiny by appropriate independent assessors (e.g., organised by the Chief Scientific Adviser).

B3. Subject to B2, all AI projects must be able to allow a third-party to: (1) investigate the algorithmic workings, use scenarios, and underlying data from an 'adversarial perspective' [5]; This might require the supplier to provide 'expert' witness/evidence of the tools' operation. All third parties will have appropriate data protection and information security policies in place.

Principle C. Explainable: The ability for any AI to provide an 'explanation' of its output will be a determining factor in its implementation.

C1. The level of explanation expected will be determined by (1) the function it performs (e.g., is it informing a high-impact decision about an individual); (2) the outputs required of it (i.e., who needs to understand what regarding the output and how was this reached).

Principle D. Responsible: All AI that affects the public will have responsible usage policies (i.e., intensions are defined before deployment so that outcomes and impact can be tracked) and procedures to ensure that users do not accept AI outputs uncritically.

D1. The ability of AI to make decisions without a human being part of that decision will be determined by the function that the AI performs.

D2. All AI that effects the public must have a human as the ultimate decision-maker.

D3. All AI will have a human or automatic means of being stopped if it displays unintended or undesired outputs.

D4. Those responsible for AI-enabled systems must proactively mitigate the risk of unintended biases or harms, during initial rollout and as they learn, change, or are redeployed.

Principle E. Accountable: All AI will have a clearly identified individual accountable for its operation and outputs.

E1. All Accountable persons and end-users will be suitably trained in the use of the relevant AI.

E2. The use of AI in CoL will be subject to proper governance and oversight at the relevant organisational level.

E3. AI enabled data sets and technology systems will be governed and assured under the same frameworks as wider data processing responsibilities, linking what is used and how it is used to the appropriate Information Asset Register and ROPA.

Principle F. Robust: All data used to train, or that is analysed by, an AI will be robust and reliable enough for its intended purpose. This requires assessing, tracking and reporting on the quality of data, by way of recognising that the quality of data dictates the quality of the analysis.

F1. All AI in CoL will be used only for the purpose it was designed, trained and authorised for.

F2. With regards to data usage, all data used in CoL AI will be subject to a Framework outlined by a governance board to guard against issues such as bias, unintended proxies, non- representativeness, unfairness, and untimeliness.

F3. the Government Office for Artificial Intelligence's Guidelines for AI procurement must inform contract implementation and management.

Principle G. Ethical: All data used to train, or that is analysed by, an AI must be ethically sourced and its lineage fully understood. This requires interrogative analysis of the supply chain to ensure that there is no risk of supporting modern day slavery and/or exploitation of the vulnerable, and where identified, ensure that we highlight and seek to undertake remedial action, including reporting to the appropriate authorities. The following ethical principles support this approach:

G1. Human agency remains paramount. We believe AI systems should support human agency and the fundamental rights of humans, and must not decrease or limit ultimate human autonomy.

G2. AI must deliver positive societal impact. AI will only be used where societal impact from its use is positive, progressive and respectful and that AI will not be used to discriminate against communities.

G3. Harm and acting within the law. Where any of our services use AI, they will do so within the law. All use will be monitored and we will guard against discriminatory bias developing within systems.

G4. Accountability. All senior Corporation officers will receive training on AI, where and

how it could be used across the organisation, and the risks associated in AI use, particularly around privacy and bias. We will be fully accountable for any decisions that an AI system or AI implementation makes.

G5. Transparency. We understand that decisions resulting from AI processes may not always be replicable by humans, but where a decision delivered by AI requires justification, the AI process used should be explainable and transparent, and the decision capable of being made by a human.

G6. Governance & Oversight. All AI will be subject to a governance and oversight process. All internally developed and externally acquired AI will be verified for on-going compliance with these principles.