Parliamentary and Health Service Ombudsman

AI

Governance Framework V1

1 Introduction

1.1 This AI governance framework is designed to enable PHSO to benefit from AI driven technology innovations whilst avoiding the very real risks. As the complaint handler of last resort for the NHS in England and government services, we need to be sensitive to the needs of the people we work for and with, some of whom are distressed or vulnerable. This means that we need, above all, to be human.

1.2 Being human is about informed use of artificial intelligence to augment how we work, to better inform our decision making and to enable us to make the best use of our people's time and expertise. Being human is about making sure that our complainants, MPs and other have online, immediate access to the information they need on where their complaints are up to or on issues in their constituencies whilst making it easier to talk to a person when that's appropriate or needed.

1.3 This framework describes how we intend the balance our human and machine capabilities, ensuring that we get the best out of both.

2 Key principles

2.1 **AI will not restrict accessibility or access to justice** | AI is great for automating at massive scale, enabling organisations to do much more at greater velocity than before. The downside is that all AI solutions are only as good as the data and logic behind them and can fall foul of bias, prejudice and hallucinations. Whilst an individual may have unconscious biases than can impact on a decision or action, that is limited to their sphere of influence. An AI making bad decisions has the potential to have a much greater sphere of influence, which for PHSO could restrict access to justice. Therefore, we will not use AI to determine conditions for access and will never automate casework decisions.

2.2 **Human in the loop** | there is always a person involved in all AI uses at PHSO processes. This means that we are never using automated processing to make decisions and remain sensitive to nuances that an AI may not pick up. This provides has an opportunity to course correct or change the outcome of the model or the AI driven process.

Examples of this at PHSO could include:

- PHSO deploys a tool to auto-redaction of personal or otherwise sensitive information for publication or release under UK: GDPR. This tool flags potentially sensitive information to a human reviewer who confirms that the machine has correctly identified information that should not be published and

if there is any further information to be withheld in the document. The machine doesn't make the decision to publish, the human does.

- PHSO uses an AI tool to match invoices to a purchase order, removing the need for manual matching. However, a finance professional validates the proposed payment schedule before transferring funds.

2.3 **Transparency** | PHSO publishes clear, and accessible information on our use of AI, how we manage and monitor accuracy and how we ensure that we do not adopt automated processing by stealth remaining consistent with the spirit and the letter of the UK's GDPR legislation. Everyone we would work with and for, should be aware when they are interacting with an artificial intelligence and not a human being.

2.4 **Information Rights**: Before deploying any AI solutions that uses personal information, PHSO will set out how we intend to uphold the rights and freedoms of individuals including instructions to cease or restrict processing.

2.5 **No automated processing|** This extends 2.1 to decisions made that could impact individual data subjects, for example, recruitment, performance reviews or identification of individuals taking part in engagement activities. Under Article 22 UK GDPR, PHSO can only carry out automated processing i.e. solely automated decision-making that has legal or similarly significant effects on data subject if the decisions made are:

a. necessary for the entry into or performance of a contract; or
b. authorised by domestic law applicable to the controller; or
c. based on the individual's explicit consent.

PHSO is unlikely to be able to meet these requirements as we do not operate under explicit consent. Whilst we collect general consent from complainants to ensure we satisfy the common law duty of confidentiality, our legal basis under GDPR is a task carried out in the public interest set out in our legislation.

2.6 **We expect our suppliers to operate to the same standards we do|** We understand what is happening and can communicate that openly and accessibly. This applies to our suppliers and their suppliers as well – we need to understand the risks posed by their deployment of AI tools and solutions and will include monitoring within our supply chain security framework.

2.7 **AI is fair and accurate** | Deployment is not the end but the beginning, need to make sure that all AI solutions are subject to rigorous and appropriate testing and controls

to ensure that fairness continues throughout the lifecycle of the AI driven application or tool.

2.8 **Accountability** | AI deployments shall be appropriately risk assessed and approved before deployment.

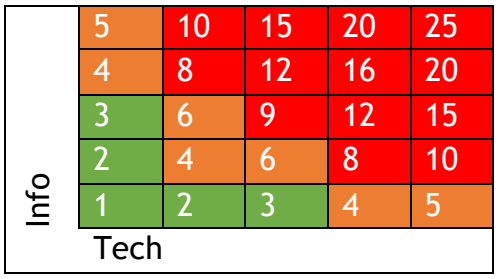
2.9 **Technical Robustness and safety** | AI systems need to be resilient and secure. They need to be safe, ensuring a fall-back plan in case something goes wrong, as well as being accurate, reliable and reproducible. That is the only way to ensure that also unintentional harm can be minimized and prevented.

2.10 **Sustainable |** AI is energy and resource intensive. As part of evaluating potential AI solutions, PHSO will consider the environmental impact as part of our commitment to sustainability.

AI Assessment Framework

2.11 This AI Assessment Framework takes a familiar risk-based approach, considering both the information that will be processed as well as the technology and the company behind it.

Traffic lights approach [Information] x [Tech] = AI risk score



2.12 The **AI Risk score** will be either green, amber or red.

Green	Ok to proceed, follow route 1
Amber	Proceed with assurances, follow route 2
Red	Do not proceed

There are two routes to approval for both green (low risk) and amber (mid risk) uses of AI. These are set out in checklist format below. All must be completed before implementation.

	Governance Route 1	Governance Route 2
DPIA Frequency of review	☐ Basic (as a minimum) ☐ Annual	☐ Full ☐ quarterly
Supplier Assurance	☐ PHSO's supply chain security standards	☐ PHSO's supply chain security standards In depth OSINT review
Design	☐ High level design to Technical Design Authority	☐ Detailed design to Technical Design Authority
Procurement and Legal risk assessment	☐ No	☐ Yes
Post implementation review and assurance	☐ Required	☐ Required
Approval	☐ TDA	☐ SIRG

Technology	
1	• Clearly understood tool, transparent processing, simple to test and maintain. UK based data processing • UK registered DPO. • Meets PHSO's supply chain security standards. • Human in the loop • Green Supplier Intelligence report
2	• Clearly understood tool, transparent processing, simple to test and maintain. • Data processed in non-UK but EU/EEA or other countries who are deemed adequate with UK GDPR. • Meets PHSO's supply chains security standards. • UK registered DPO / good score supplier scorecard • Human in the loop • Green Supplier Intelligence report
3	• Processed outside of UK/EEA/GDPR adequate countries • Partially meets PHSO's supply chains security standards. • No feedback mechanism • Human in the loop • Green Supplier Intelligence report
4	• Processed outside of UK/EU/GDPR adequate countries • Does not meet PHSO's supply chains security standards. • More than one data breach last 24 months • ICO notice/action • No feedback mechanism • No human in the loop • Amber Supplier Intelligence report
5	• Opaque, no transparency, no testing and control mechanism • Does not meet PHSO's supply chains security standards. • Processed outside of UK/EU/GDPR adequate countries • More than two data breaches last 24 months • ICO notice/action • No feedback mechanism • No human in the loop • Red Supplier Intelligence report

Information	
1	The information is low risk. It does **not** contain: • personal information • complaint information and material evidence • information that could be commercially sensitive • protected legal advice.
2	The information is still low risk, but may contain: • limited personal, financial or complaint information. It does not contain: • special category data • material evidence obtained during our investigations.
3	The information may contain: • Personal information i.e. information that identifies or relates to an individual • Information that whilst not personal or identifying, could be put together with external information to identity individuals.

4	• The information may cause harm or distress to data subjects if released or includes:
	  ○ Special category data
	  ○ Commercially sensitive data
	  ○ Legal advice
	  ○ Potentially damaging information

5	The information will cause harm or distress to data subjects if released/combined in mosaic effect, significant reputational risk Special category data Evidence/material evidence

Version	Summary	Date Approved	Approved by:
0.1	Draft	3 October 2024	Technical Design Authority
1	Initial AI Governance Framework, sent for review across multiple disciplines	7 November 2024	SIRO