REFORM

GETTING THE MACHINE LEARNING: APPENDIX 3

Freedom of Information Act requests and responses

September 2024

1. Methodology

Six separate FOI requests were sent in June 2024. Four of these requests were sent to the following 16 Government Departments and NHS England (NSHE).

- Cabinet Office (CO)
- Department for Business and Trade (DBT)
- Department for Culture, Media and Sport (DCMS)
- Department for Education (DfE)
- Department for Energy Security and Net Zero (DESNZ)
- Department for Environment, Food & Rural Affairs (DEFRA)
- Department for Science, Innovation and Technology (DSIT)
- Department for Transport (DfT)
- Department for Work and Pensions (DWP)
- Department of Health and Social Care (DHSC)
- Foreign, Commonwealth & Development Office (FCDO)
- HM Revenue & Customs (HMRC)
- HM Treasury (HMT)
- Home Office (HO)
- Ministry of Housing, Communities and Local Government (MHCLG) (formerly the Department for Levelling Up, Housing and Communities (DLUHC))
- Ministry of Defence (MoD)
- Ministry of Justice (MoJ)

One request was sent only to DBT, DEFRA, DWP, HMRC, HO and MoJ.

One request was sent only to the Cabinet Office.

Analysis is based on responses received up until 1 August 2024.

Several questions received few responses, with the majority denied under one of the exemptions to the Freedom of Information Act. As a result, they were not used in the research paper. This Appendix summarises the responses to the questions where there were several useful responses. The exemptions used by departments to questions are listed below:

- Section 12 of the Freedom of Information Act (FOIA) permits public authorities to not comply with an FOI request if the cost of complying would exceed the appropriate limit. For central government departments the appropriate limit is £600.
- Section 21 of the FOIA permits public authorities to not comply with an FOI request if the information requested is reasonably accessible to the applicant.
- Section 24 of the FOIA permits public authorities to not comply with an FOI request for the purposes of safeguarding national security.
- Section 31 of the FOIA permits public authorities to not comply with an FOI request if the information requested would, or would be likely to, prejudice the prevention or detection of a crime.
- Section 35 of the FOIA permits public authorities to not comply with an FOI request if the information requested related to the formulation or development of Government policy.

2. FOI 1

2.1 Request

The following FOI request was sent to the CO, DBT, DCMS, DfE, DESNZ, DEFRA, DLUHC, DSIT, DfT, DWP, DHSC, FCDO, HMT, HMRC, HO, MoD, MoJ and NHSE in June 2024.

1) Any template Data Protection Impact Assessments and Data Sharing Agreements used in your department.

2) Any template business case documents used for approving technology projects in your department.

3) Any guidance issued on the use of 'coding co-pilots' and other kinds of AI to develop software in your department.

4) Any internal guidance or staff policies on the procurement of AI specific to your department.

5) Any technical guidance on using AI software specific to your department, for example bias checking, model monitoring practices or red teaming processes.

2.2 Responses

Department	Question one	Question two	Question three	Question four	Question five
CO	No response	No response	No response	No response	No response
DBT	Yes	Yes	Yes	Yes	Yes
DCMS	No response	No response	No response	No response	No response
DfE	Yes	Yes	Yes	Yes	Yes
DEFRA	Yes	Yes	Yes	Does not hold this information	Does not hold this information
DESNZ	No response	No response	No response	No response	No response
DLUHC	No response	No response	No response	No response	No response
DSIT	Yes	Yes	Yes	Yes	Yes
DfT	Yes	Yes	Yes	Yes	Yes
DWP	Yes	Section 21	Yes	Section 21	Yes
DHSC	Yes	Yes	Does not hold this information	Does not hold this information	Section 35
FCDO	No response	No response	No response	No response	No response
HMRC	Yes	Yes	Yes	Yes	Yes
HMT	Section 12	Section 12	Section 12	Section 12	Section 12
HO	No response	No response	No response	No response	No response
MoD	Requested clarification	Requested clarification	Requested clarification	Requested clarification	Requested clarification
MoJ	Yes	Yes	Yes	Yes	Yes
NHSE	Requested clarification	Requested clarification	Requested clarification	Requested clarification	Requested clarification

2.3 Results

“Any template Data Protection Impact Assessments and Data Sharing Agreements used in your department.”

DBT provided the following documents:

- ‘Data Protection Impact Assessment (DPIA): Stage One: Screening (Generative AI Specific Use Case Version)’
- ‘Data Protection Impact Assessment (DPIA): Stage Two: Data Processing Information — Generative AI (Specific Use Case) Version’
- ‘Data Protection Impact Assessment (DPIA): Stage Two: Data Processing Information — Generative AI Models (LLMS)’
- ‘Data Protection Impact Assessment (DPIA): Stage Two: Data Processing Information’

- ‘Data Protection Impact Assessment (DPIA): Stage Three: Risk Assessment’

DfE provided the following documents:

- ‘Independent Controller to Independent Controller Agreement for Making: (1) DfE Data Extracts, and/or (2) Matched Datasets available in Secure Environment(s)’
- ‘Independent Controller to Independent Controller Agreement for the direct supply of: (1) DfE Data Extracts; and/or (2) Matched Datasets’
- ‘Individual Declaration Form: Making the DfE Data Extract and/or Matched Dataset available in Secure Environment(s)’
- ‘Individual Declaration Form: Supply of DfE Data Extracts and/or Matched Datasets’
- ‘Individual Declaration Form for Joint Requester Permitter Users: Supply of DfE Data Extracts and/or Matched Datasets’
- ‘Data Protection Impact Assessment (DPIA) form’

DEFRA provided the following documents:

- ‘Data Protection Impact Assessment (DPIA)’
- ‘Controller to Controller Data Sharing Agreement Guidance’

DSIT provided the following documents:

- ‘Data Protection Impact Assessment – (Name of project)’
- ‘Data Sharing Agreement’

DfT provided the following documents:

- ‘Data Protection Impact Assessment (DPIA)’
- ‘Model data sharing memorandum of understanding’

DWP provided the following documents:

- ‘DWP Data Protection Impact Assessment Part 2 – detailed assessment’
- ‘Data Sharing Agreement’

DHSC provided the following documents:

- ‘Data Protection Impact Assessment (DPIA)’
- ‘Data Sharing Agreement’

HMRC provided the following documents:

- ‘UK GDPR and DPA 2018 Compliance Memorandum of Understanding (MoU) on Data Protection (Intra-Crown Controller to Processor)’
- ‘Memorandum of Understanding (Process)’

MoJ provided the following documents:

- ‘Stage 1 DPIA Screening’
- Stage 2 Full DPIA’
- ‘Draft Data Sharing Agreement’

These documents are available from *Reform* upon request.

“Any guidance issued on the use of 'coding co-pilots' and other kinds of AI to develop software in your department.”

DfE, DEFRA and DfT’s said they said that they follow CDDO guidance in this area.

DSIT said that they follow CDDO guidance and “we do not have specific departmental recommendations, we ask teams to review best practice documentation provided by GitHub Copilot Trust centre”.

DWP, HMRC and MoJ said they are not currently using coding co-pilots.

- DWP: “DWP are not currently using any AI to develop software in the Department”.
- HMRC: “Artificial Intelligence (AI) coding assistants are not currently being used by HMRC in a production capacity, although small scale testing has been done to build our understanding of how this capability might be deployed in future. Should we deploy this capability in future we will ensure the appropriate controls and guidance is in place. Where we use AI in a way that could impact customer outcomes, we always ensure that the result is explainable, that there’s a human in the loop, and that it complies with our data protection, security, and AI ethics standards”.
- MoJ: “We currently do not use any generative AI based assistants to help us develop any software”.

DBT said that they do not allow the use of coding co-pilots at all: “we explicitly do not allow any ‘copilot’ for coding, software procurement or access at this time”.

“Any technical guidance on using AI software specific to your department, for example bias checking, model monitoring practices or red teaming processes.”

DfE and DfT said that they follow CDDO guidance in this area.

DSIT said that they follow CDDO guidance and “are working with colleagues across the department to ensure any DSIT AI specific policy is fit for purpose including considerations of risks and burden”.

DBT and DWP said that they have no technical guidance.

- DBT: “We are not aware of any technical guidance surrounding specific tools. The Data Science team occasionally provide additional technical support where required, but this consists of a data scientist consulting with the team who wishes to deploy a tool rather than the availability of documentation”.
- DWP: “There is no technical guidance that is specific to this department”.

HMRC said that “safe and ethical use of AI is paramount for HMRC. We have an established AI Assurance process, AI Ethics Framework, and governance to ensure effective, responsible use of all AI models deployed within HMRC that is aligned with government wide guidance… We are working to develop our existing framework as our understanding and use of AI evolves and matures. As the regulatory framework for use of AI in government develops further, we will adapt our framework accordingly, alongside utilising our knowledge and experience to engage and help shape and influence the strategic direction.”

MoJ said that “in addition to our standard digital design, and delivery processes, we are currently designing assurance checks, for any projects that use AI. These checks are based on the Turing Institute ethics and safety guidelines (FAST principles) along with the Government generative AI framework. Our Cyber red teaming is an independent activity.”

3. FOI 2

3.1 Request

The following FOI request was sent to the CO, DBT, DCMS, DfE, DESNZ, DEFRA, DLUHC, DSIT, DfT, DWP, DHSC, FCDO, HMRC, HMT, HO, MoD, MoJ and NHSE in June 2024.

1) Details of the use cases reported by the department to the NAO in their report 'Use of artificial intelligence in government', including the state of deployment and the time those were shared.

2) The department's DevOps policy, Software release processes, and policies about deployment into production environments.

3.2 Responses

Department	Question one	Question two
CO	Yes	Yes
DBT	No response	No response
DCMS	Yes	Section 31
DfE	Yes	Yes
DEFRA	No response	No response
DESNZ	No response	No response
DLUHC	Requested clarification	Requested clarification
DSIT	No response	No response
DfT	Yes	Section 24
DWP	Section 31	Section 31
DHSC	Yes	Yes
FCDO	No response	No response
HMRC	Yes	Section 31
HMT	Yes	Does not hold this information
HO	No response	No response
MoD	Section 21	Section 21
MoJ	Yes	Yes
NHSE	Yes	Section 21

3.3 Results

“Details of the use cases reported by the department to the NAO in their report 'Use of artificial intelligence in government', including the state of deployment and the time those were shared.”

DfT, DHSC, HMRC, HMT, MoJ and NHSE responded with details of the use cases reported to the NAO.

Figure 1: DfT’s use cases response

“The Department submitted three use cases; these are listed below:

1) Using AI to improve our response time and accuracy in relation to public consultation responses.
2) Exploring using image analysis to detect fraud, particularly around the Electric Vehicle Homecharge Scheme (EVHS).
3) Exploring using image recognition to support audits or whether vehicles are effectively taxed.

All these use cases are in the pilot and concept stage.”

Figure 2: DHSC's use cases response

"We provided one use case to the National Audit Office (NAO) in December 2023. This was as follows:

Use case 1: First Alert (from Dataminr)

- Status: fully deployed.
- AI use case was procured commercially 'off the shelf'.
- Purpose was to support operational decision-making such as prioritisation, eligibility, and enforcement and to support research or monitoring. In the event of a major incident the tool should improve response times and better inform immediate decisions. It is also useful for longer-term monitoring, allowing a story to be tracked across many information sources much more efficiently than if done manually.
- Expected impacts were to:
    - Improve service speed
    - Improve service quality
    - Service cost reduction (including freeing up resources for other tasks)

As well as this specific use case, we also returned information to the NAO on planned and piloted use cases. We noted there were approximately five AI use cases being planned or piloted and the information provided to the NAO was as follows (noting these were given as examples and not yet deployed):

1) We already identify trends in text data using topic modelling, but those topics are manually interpreted. We are piloting use of BERT to automate the topic labelling. The outcome is an ability to fully productionise analyses that currently can't be fully automated to improve consistency, save resources, and facilitate real-time automated feeds.

2) Performance metrics that focus on a single measure reward gaming of that measure giving perverse outcomes. We are piloting approaches whereby a more holistic set of data is fed into a model to allow comparative performance to be assessed for a similar set of hypothetical inputs to the trained model. The outcome could be a more holistic performance measure that avoids measurement cliff-edges and ability to game performance metrics, ultimately giving more freedom to providers to take a broader view of good performance too rather than focussing on meeting a narrow evaluation."

Figure 3: HMRC's use cases response

"HMRC provided a questionnaire response in September 2023. This included some examples of existing AI use cases that are already in production:

- An analytics and debt management system featuring decision engine capabilities.
- Predictive analytics (using decision trees) to support the identification of non-compliance in the tax system, including tax credit error and fraud.
- HMRC's digital assistant that automatically helps customers to find the information they are looking for."

Figure 4: HMT's use cases response

"HM Treasury provided information on three use cases to the NAO in November 2023.

- Triage Correspondence – using AI to predict the right team to allocate inbound correspondence. Now live, in pilot at time of submission.
- Information Assistants – using OpenAI's GPT model to help us assess and review research materials. No longer active, in planning at time of submission.
- Talk to your data – a semantic layer on top of data in the warehouse and allowing the natural language generation of questions and answers. No longer active, in planning at time of submission."

Figure 5: MoJ's use cases response

"The NAO did not list the specific details of these use cases. However please find the information below:

- Short term custody predictor – no generative AI used.
- Actuarial Risk Assessment Instruments (ARIAs) – no generative AI used.
- Short Term Prison Demand Modelling – no generative AI used.

All were fully deployed at the time this was shared with the NAO on the 23 November 2023."

Figure 6: NHSE's use cases response

"The use cases are as follows:

Use case 1: Auto moderation of user reviews left on nhs.uk

- Status: AI components currently running in production system in 'shadow parallel run' alongside human moderators, pending migration in early 2024.
- Confirmed this used generative AI. NB while use of generative AI tools was tested for use in production model, ultimately it did not reach performance thresholds and was not deployed. Generative AI has been used to create training data on which more traditional NLP models were developed.
- AI use case was developed in-house with internal expertise
- Purpose was to replace manual moderation of comments, delivering efficiencies (but no direct implications for service /care delivery). Automation of comment moderation allows for faster publication of reviews with better user experience, delivers efficiencies, allows for better scaling / marketing of the reviews component of nhs.uk by reducing impact of manual moderation bottleneck.
- Expected impacts were to:
    - Improve service speed
    - Improve service quality
    - Service cost reduction (including freeing up resources for other tasks)
    - Development of new services"

Figure 7: NHSE's use cases response (continued)

Use case 2: A&E Admissions Forecasting Tool hosted on National Data Platform (Foundry)

- Fully deployed
- Not using generative AI
- AI use case was developed in collaboration with commercial suppliers or other non-public sector partners.
- Purpose was to support operational decision-making such as prioritisation, eligibility, and enforcement. The A&E admissions tool provides a three-week forecast of expected admissions via A&E to support operational planning. Forecasts are made available to users working in local health systems to national roles.
- Expected impacts were to:
    - Improve service speed
    - Improve service quality
    - Service cost reduction (including freeing up resources for other tasks)
    - Planning resource allocation, staffing and beds

Use case 3: COVID-19 Early Warning System hosted on National Data Platform (Foundry)

- Was fully deployed but had been mothballed and is no longer a live operational product providing forecasts to users, however it still produces a limited output for internal monitoring only to ensure that it can be scaled up if needed in future.
- Not using generative AI
- AI use case was developed in collaboration with commercial suppliers or other non-public sector partners.
- Purpose was to support operational decision-making such as prioritisation, eligibility, and enforcement. The COVID-19 EWS provided forecasts of expected hospital admissions and beds use for COVID-19 patients.
- Expected impacts were to:
    - Improve service speed
    - Improve service quality
    - Service cost reduction (including freeing up resources for other tasks)
    - Supporting pandemic response, planning resource allocation

As well as those specific use cases which had been deployed in some form, we also returned information to the NAO on planned and piloted use cases. We noted there were approximately 7 AI use cases currently being planned or piloted and shared the following as examples:

- ChatGPT Implementation for Model Health System
    - Description: Implementing ChatGPT to answer common user queries within the Model Health System.
    - Pilot Status: Undergoing real-world testing and feedback collection.
- Interactive Analytics with Chat GPT
    - Description: Enabling data-analyses and responses through the Model Health System, using ChatGPT.

Figure 8: NHSE's use cases response (continued)

"As well as those specific use cases which had been deployed in some form, we also returned information to the NAO on planned and piloted use cases. We noted there were approximately 7 AI use cases currently being planned or piloted and shared the following as examples:

- ChatGPT Implementation for Model Health System
  - Description: Implementing ChatGPT to answer common user queries within the Model Health System.
  - Pilot Status: Undergoing real-world testing and feedback collection.
- Interactive Analytics with Chat GPT
  - Description: Enabling data-analyses and responses through the Model Health System, using ChatGPT.
- Custom GPT for Clinical Coding
  - Description: A custom GPT model for automating clinical coding from notes.
  - Pilot Status: Demonstrating its coding capabilities. Has passed the coding exam.

4. FOI 3

4.1 Request

The following FOI request was sent to the DSIT in June 2024.

1) The responsibilities of the AI Directors' Policy Board.

4.3 Results

"The responsibilities of the AI Directors' Policy Board."

Figure 9: The AI Directors' Policy Board

"The primary objectives of the UK AI Directors' Policy Governance Board are to:

- Policy Development: Inform the board on DSIT-led cross cutting AI policies and seek consultation from wider government departments.
- Risk Management: Assess the impact of mitigations and whether the residual risk is tolerable, assessing departments' confidence in their risk returns and where coordination / join-up is required.
- Transparency and co-ordination: Engage relevant directors on a routine basis to share information on AI-related departmental work-streams, address concerns, gather feedback on AI initiatives and co-ordinate as required.

ABOUT REFORM

Reform is established as the leading Westminster think tank for public service reform. We believe that the State has a fundamental role to play in enabling individuals, families and communities to thrive. But our vision is one in which the State delivers only the services that it is best placed to deliver, within sound public finances, and where both decision-making and delivery is devolved to the most appropriate level. We are committed to driving systemic change that will deliver better outcomes for all.

We are determinedly independent and strictly non-party in our approach. This is reflected in our cross-party Advisory Board and our events programme which seeks to convene likeminded reformers from across the political spectrum.

Reform is a registered charity, the *Reform* Research Trust, charity no. 1103739.

REFORM

@reformthinktank

info@reform.uk

www.reform.uk