

OpenAI
Pioneer Building
3180 18th Street
San Francisco
USA

By email: [REDACTED] ([REDACTED]@openai.com)

15 June 2023

Dear [REDACTED],

I write further to OpenAI's 14 April response to the ICO's blogpost, 'Generative AI: eight questions that developers and users need to ask'.

As part of our ongoing work to monitor data protection compliance among organisations developing and deploying generative AI services, we are requesting the following information:

1. Copies of any Data Protection Impact Assessments (DPIA) that you have prepared in relation to your Generative Pretrained Transformer (GPT) large language models (LLMs) (as referenced in your 14 April response). If any of the information requested in points 2 to 10 is included in DPIAs, please provide references to the relevant section(s).
2. In respect of any personal data processed on the basis of legitimate interests (as referenced in your 14 April response), copies of any Legitimate Interests Assessments (LIAs) undertaken by OpenAI. If LIAs have not been produced, please describe how you conducted and documented the requisite purpose, necessity, and balancing test for any processing which relies upon legitimate interests as a basis for processing.
3. A complete list of the open-source datasets used to train your models and any assessments of the extent to which personal data is contained within these datasets.
4. A complete list of any third parties and sources from which you have obtained personal data to train your models.

5. A description of any re-identification testing (eg intruder testing) you carry out to assess the risk of personal data being obtained from model outputs.
6. A description of any output filtering you perform to identify and remove personal data from model outputs.
7. A description of steps taken to ensure personal data used to train your models is accurate and steps taken to ensure inaccurate personal data that is part of the training data is erased or rectified without delay, in line with the accuracy principle of UK data protection law.
8. A description of any technical and organisational measures taken to mitigate the risks of security attacks targeted at your models, including:
 - a. Model inversion;
 - b. Membership inference; and
 - c. Data leakages in the model's outputs
9. Confirmation of the number of data protection requests or complaints received by OpenAI from UK data subjects since 1 January 2023. Please also include a summary of the overarching issues covered by any requests or complaints.
10. Details of any assessment of the likelihood of users under the age of 18 accessing your services and, where required, any assessment of the services' conformance with the ICO's Children's code.

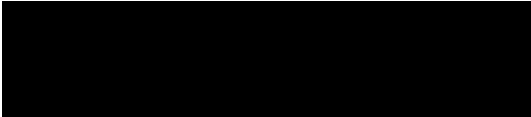
I would be grateful if you could provide the requested information by 6 July 2023.

For awareness, the Commissioner may consider the use of an Information Notice, as per Section 142 of the Data Protection Act 2018 (DPA18), in the event that OpenAI fails to provide the requested information. Failure to comply with an Information Notice is an offence and can result in:

- A fine, and / or
- An Information Order through the courts, as per Section 145 of the DPA18.

Should you wish to discuss this request, please do not hesitate to contact me.

Yours sincerely,



Stephen Almond

Executive Director – Regulatory Risk

Information Commissioner's Office

Please note that we are often asked for copies of the correspondence we exchange with third parties. We are subject to all of the laws we deal with, including the United Kingdom General Data Protection Regulation, the Data Protection Act 2018 and the Freedom of Information Act 2000. You can read about these on our website (www.ico.org.uk).

Please say whether you consider any of the information you send us is confidential. You should also say why so that we can take that into consideration. However, please note that we will only withhold information where there is good reason to do so.

For information about what we do with personal data see our privacy notice at www.ico.org.uk/privacy-notice