Cursory assessment of OpenAI's Chat GPT, DALL-E et al. 2023-05-08

Background

The licencing function of KTH VS/IT have received a number of purchase requests where researchers and other employees of the government agency want to buy licenses for Chat GPT, DALL-E and other services from OpenAI.

IT-SMC have made a cursory assessment to be able to give guidance.

Documentation used for the review

Agreements and terms:

- https://openai.com/policies/terms-of-use
- https://openai.com/policies/privacy-policy
- https://openai.com/policies/data-processing-addendum
- record-Data Processing Agreement (KTH and OpenAI).pdf
  - An automatically generated agreement proposal from their agreement service, where we aborted the process before we entered into the agreement
- OpenAI Information Security Addendum in DPA - Exhibit B ANNEX II
  - We sent a request to legal@openai.com in order to get access to the document and they sent OpenAI Information Security Addendum (March 2023).pdf
- https://platform.openai.com/subprocessors

How OpenAI use data that is processed in the service:

- https://platform.openai.com/docs/data-usage-policies
- https://help.openai.com/en/articles/7039943-data-usage-for-consumer-services-faq
- https://help.openai.com/en/articles/5722486-how-your-data-is-used-to-improve-model-performance

OpenAI in Microsoft Azure:

- https://learn.microsoft.com/en-us/legal/cognitive-services/openai/limited-access
- https://azure.microsoft.com/en-us/products/cognitive-services/openai-service/#overview
- https://learn.microsoft.com/en-us/azure/cognitive-services/openai/overview#comparing-azure-openai-and-openai
- https://learn.microsoft.com/en-us/azure/cognitive-services/openai/quickstart?tabs=command-line&pivots=programming-language-studio
- https://customervoice.microsoft.com/Pages/ResponsePage.aspx?id=v4j5cvGGr0GRqy180BHbR7en2Ais5pxKtso_Pz4b1_xUOFA5Qk1UWDRBMjg0WFhPMkIzTzhKQ1dWNyQlQCN0PWcu
- https://learn.microsoft.com/en-us/azure/cognitive-services/openai/faq
- https://learn.microsoft.com/en-us/legal/cognitive-services/openai/data-privacy
- https://customervoice.microsoft.com/Pages/ResponsePage.aspx?id=v4j5cvGGr0GRqy180BHbR7en2Ais5pxKtso_Pz4b1_xURE01NDY1OUhBRzQ3MkQxMUhZSE1ZUlJKTiQlQCN0PWcu

Assessment

OpenAI divides their offering in Consumer services and Commercial services with significantly different approaches to what applies to the data processed in them.

Data Processing Agreement and processing of personal information

Consumer services (ChatGPT, DALL-E)	Commercial services (API)
It is not possible to enter into a Data Processing Agreement with EU's standard contractual clauses for what OpenAI presents as their consumer services: " (ChatGPT, DALL-E Labs)". That means that KTH cannot process personal information in those services without violating the Data Protection Act.	KTH can enter into a Data Processing Agreement with EU's standard contractual clauses for what OpenAI's presents as their commercial services: "... business service offerings (APIs for text completion, images, embeddings, moderations, etc.)". Please note that it is KTH's delegation of authority that controls who has the right to enter into an agreement on KTH's behalf and that the agreements needs to be registered. All processing is done in a third country (USA). If KTH signs their Data Processing Agreement we accept their current sub processors. They will inform about changes regarding the sub processors they enlist, and we can end the agreement if we do not accept the changes. There are paragraphs in the agreement that indicates that they might try to address possible objections, but our assessment is that it is unlikely that they will to that to cater for individual customers interests. Please note that OpenAI stipulate several requirements on KTH as a customer under " 6. Obligations of Customer.", e.g.: • KTH's responsibility includes, but is not limited to, keeping a record of "completion ID" or similar ID number related to specific data subjects in order to assist OpenAI when they handle individuals that want to exercise their rights under the GDPR. • KTH shall ascertain compliance with the Data Protection Act and other laws regarding data protection. Also note that the Data Processing Agreement explicitly states under B. DESCRIPTION OF TRANSFER that no sensitive data is intended to

	be transferred unless the user includes it unexpectedly in unstructured data.

Confidentiality agreement

OpenAI refers in several of their documents to that their personnel as well as the personnel of their sub processors are subject to commercial, contractually regulated confidentiality and special security regulations.

Use of data for training

Consumer services (ChatGPT, DALL-E)	Commercial services (API)
They train their models on customer data per default in their consumer services. It is possible to opt out via a web form. OpenAI also describes 4 scenarios where authorized personnel with their company access the data that is processed in the service: 1) security and abuse incidents, 2) support requests, 3) to comply with legal requests and 4) for model training (if we have not opted out).	OpenAI changed the practices for their commercial API services starting 1st of March 2023 and no longer per default use data customers expose to them via the API to train OpenAI's models. They store data for 30 days in order to handle abuse incidents and misuse. A small, authorized group of personnel with OpenAI and their sub processor may access that data while reviewing such cases. They note the possibility of making a separate agreement to exempt enterprise customers from this practice under certain conditions where the risk of abuse or misuse is low. If we opt in for model training for API use, it is probable that certain OpenAI personnel will have access to the data. Our assessment is that they also have access to the data to comply with legal demands and for support requests.

Intellectual property rights, sensitive data, ownership

Data transferred to OpenAI is processed in a third country and might be handled by personnel with OpenAI or their sub processors.

According to OpenAI's Terms of Use, 3. Content (a) Your Content, KTH owns its "Input" and have the rights to its "Output". KTH at the same time accepts that OpenAI may use "Content" ("Input" and Output" collectively) to provide and maintain the Services, comply with applicable law, and enforce their policies.

Note regarding Azure OpenAI Service

While assessing OpenAI's services we noted that Azure OpenAI Service provides API access to OpenAI's Language Models in Azure with the option to process the information within EU/EES. At present, the customer is required to apply for access to Azure OpenAI Service for their proposed usage scenario.

Azure OpenAI Services do not train their models on the data the customer exposes through the API.

Prompts and Completions are stored for 30 days and monitored by Microsoft's personnel to detect usage that conflicts with the terms of use. It is possible to apply to be exempt from monitoring.

KTH already has a Data Processing Agreement with Microsoft that covers Azure. Commercial, contractually regulated confidentiality for Microsoft's personnel and their sub processors is part of KTH's agreement with Microsoft.

Assessment and guidance

- If personal data is to be processed, then OpenAI:s API-service is the only usable option
    - You need to perform a Data Protection Impact Assessment, https://intra.kth.se/anstallning/anstallningsvillkor/att-vara-statligt-an/behandling-av-person/konsekvensbedomningar-for-behandling-av-personuppgifter-1.1151189
    - A Data Processing Agreement must be signed
    - Avoid processing sensitive personal data
    - Note the responsibility put upon users of the API service if you process data subject's personal information in the service
- Avoid exposing sensitive and valuable information, e.g. information that might lead to patents or incur damage or economical loss for KTH if the information was exposed to others
- *If it's possible for you to get access to Azure OpenAI Services, our assessment is that it's the better option from a compliance perspective – the guidance above (for OpenAI) applies for Azure OpenAI Services as well*

If there are significant changes in the service or in how the service is used, a new assessment needs to be made. The same applies for changes in laws, regulations or statements from compliance authorities that affects the basis for using the service.

2023-05-08

För IT-SMC,
Johan Berglund
jbe@kth.se
KTH GVS/IT/IT-Stab