

Report of the work undertaken by the ChatGPT Taskforce

23 May 2024

Table of contents

- DISCLAIMER..... 3
- 1 Background 4
- 2 ONGOING INVESTIGATIONS..... 5
- 3 PRELIMINARY VIEWS..... 6
 - 3.1 Lawfulness..... 6
 - 3.1.1 Collection of training data, pre-processing of the data and training 6
 - 3.1.2 ChatGPT input, output and training..... 7
 - 3.2 Fairness 7
 - 3.3 Transparency and information obligations..... 8
 - 3.4 Data Accuracy..... 8
 - 3.5 Rights of the data subject 9
- 4 ANNEX (QUESTIONNAIRE)..... 9

DISCLAIMER

The positions presented in this document result from the coordination of the members of the ChatGPT taskforce with a view to handling investigations regarding the service ChatGPT provided by the US based company OpenAI OpCo, LLC. They reflect the common denominator agreed by the Supervisory Authorities in their interpretation of the applicable provisions of the GDPR in relation to the matters that are within the scope of their investigations. The positions presented in this document do not prejudice the analysis that will have to be made by the Supervisory Authorities in each investigation respectively. In particular, it must be taken into account that the circumstances of the investigations may change over time.

1 BACKGROUND

1. In the recent past, numerous large language models (hereinafter “LLMs”) have emerged for use in various fields.¹ While these models can offer great benefits to the public, processing operations associated with LLMs shall comply with the GDPR. It has to be noted that LLMs are trained and enhanced using a huge amount of data, including personal data.
2. Some of the most popular and widely known LLMs are those in the “GPT” category,² since it has been the first consumer-facing model to be launched on 30 November 2022 through the ChatGPT service. Several Supervisory Authorities (hereinafter “SAs”) have initiated data protection investigations pursuant to Article 58(1)(a) and (b) GDPR against OpenAI OpCo, LLC (hereinafter “OpenAI”) as controller for processing operations carried out in the context of the ChatGPT service.³
3. Until 15 February 2024, OpenAI did not have an establishment in the European Union.⁴ Insofar, as no cooperation procedures according to the One-Stop-Shop (hereinafter “OSS”) mechanism under the GDPR could apply, the European Data Protection Board (hereinafter “EDPB”) on 13 April 2023 [decided](#) to establish a taskforce to foster cooperation and exchange information on possible enforcement actions on the processing of personal data in the context of ChatGPT (hereinafter “ChatGPT TF”, SAs participating in the ChatGPT TF are hereinafter referred to as “TF members”). As the OSS did not apply, it was in particular necessary to coordinate national cases.
4. In the Plenary meeting of the EDPB on 16 January 2024, the [decision](#) was made to specify the mandate of the task force and to publish a report, outlining the interim results of the ChatGPT TF. According to this mandate, the taskforce shall:
 - Exchange information between SAs on engagement with OpenAI and on-going enforcement activities concerning ChatGPT.
 - Facilitate coordination of external communication by SAs concerning enforcement activities in the context of ChatGPT.
 - Swiftly identify a list of issues on which a common approach is needed in the context of different enforcement actions concerning ChatGPT by SAs.
5. Considering the confidential nature of the investigations, this report refers to public available information as additional source to provide information about transparency, fairness, data accuracy and data subjects’ rights towards the public.

¹ Large language models (LLMs) are deep learning models (a subset of machine learning models) that are pre-trained using vast amounts of data. Analysing these massive datasets enables the LLM to learn probability relationships and become proficient in the grammar and syntax of one or more languages. LLMs generate coherent and context-relevant language. To put it simply, LLMs respond to human language by producing coherent text that appears human-like. Most recent LLMs such as OpenAI’s GPT models are based on a neural network architecture called a transformer model.

² GPT systems are general-purpose AI models according to the definition laid down by Article 3(63) of the EU Artificial Intelligence Act (adopted, but not yet officially published as of 23 May 2024).

³ The investigations of the SAs covered the versions GTP 3.5 to GTP 4.0.

⁴ According to the “Europe privacy policy” of OpenAI, Version 15 December 2023, effective by 15 February 2024, this is OpenAI Ireland Limited.

6. As already outlined in the priorities of the EDPB for 2024-2027, providing further guidance on the interplay between the application of the GDPR and other EU legal acts, particularly the EU Artificial Intelligence Act, holds significant importance.⁵
7. Nonetheless, in line with the principle of accountability stipulated in Article 5(2) and Article 24 of the GDPR, controllers processing personal data in the context of LLMs shall take all necessary steps to ensure full compliance with the requirements of the GDPR.⁶ In particular, technical impossibility cannot be invoked to justify non-compliance with these requirements, especially considering that the principle of data protection by design set out in Article 25(1) GDPR shall be taken into account at the time of the determination of the means for processing and at the time of the processing itself.

2 ONGOING INVESTIGATIONS

8. OpenAI was found to have a single establishment in the European Union since 15 February 2024.⁷ Consequently, from that date onwards, the OSS framework applies for the “cross-border processing” carried out by OpenAI and the Lead SA within the meaning of Article 56 GDPR is responsible for exercising corrective powers where required. However, this is without prejudice to ongoing investigations of the respective SAs whose subject matter involves processing operations carried out until 15 February 2024 and that concern possible infringements of non-continuing or non-continuous nature.⁸ In this respect, these national investigations will continue to be coordinated within this TF.
9. There were several sessions of the ChatGPT TF during the reporting period. As part of the activities, a common set of questions (hereinafter, “questionnaire”) was developed, which is attached to this report as Annex. Several SAs used this questionnaire as a starting basis for their exchanges with OpenAI. The development of the questionnaire aimed to promote a coordinated approach to the investigations.
10. The privacy policies in the versions prior to 15 February 2024 are within the scope of the investigations of the respective SAs. It has to be noted that OpenAI updated their “EEA privacy policy” on 15 December 2023, effective by 15 February 2024.⁹
11. Furthermore, it has to be noted that OpenAI has already implemented a set of measures in order to comply *inter alia* with the Italian SA’s urgent decision issuing a temporary ban with reference to the ChatGPT service in Italy and the subsequent decision to lift the temporary limitation adopted on 11 April 2023.¹⁰

⁵ https://www.edpb.europa.eu/our-work-tools/our-documents/strategy-work-programme/edpb-strategy-2024-2027_en (last accessed on 23 May 2024).

⁶ Regarding the principle of accountability see para 19 of this Report.

⁷ In line with the EDPB Guidelines 8/2022 on identifying a controller or processor’s lead Supervisory Authority V 2.1, it has to be noted that the relevant lead SA is not bound to the view of a controller and can question the determination of the single or main establishment where necessary.

⁸ Regarding possible infringements of a continuing or continuous nature, the creation of a main or single establishment or its relocation from a third country to the EEA (in a procedure which was initially started without cooperation) mid-procedure allows the controller to benefit from the OSS and every pending proceeding should be transferred to the SA of the Member State in which the establishment is located. See EDPB Opinion 8/2019 on the competence of a supervisory authority in case of a change in circumstances relating to the main or single establishment, adopted on 9 July 2019, para 16, 30-32.

⁹ <https://openai.com/policies/eu-privacy-policy/> (last accessed on 23 May 2024).

¹⁰ <https://www.gpdp.it/web/guest/home/docweb/-/docweb-display/docweb/9874702#english> (last accessed on 23 May 2024).

3 PRELIMINARY VIEWS

12. The investigations conducted by the respective SAs are currently ongoing¹¹ and it is not yet possible to provide a full description of the results. Therefore, considerations in this report are to be regarded as preliminary view on certain aspects of the investigations.

3.1 Lawfulness

13. In general, it has to be recalled that each processing of personal data must meet at least one of the conditions specified in Article 6(1) and, where applicable, the additional requirements laid out in Article 9(2) GDPR.¹²
14. When assessing the lawfulness, it is useful to distinguish the different stages of the processing of personal data.¹³ In the present context, the stages can be categorised into i) collection of training data (including the use of web scraping data or reuse of datasets),¹⁴ ii) pre-processing of the data (including filtering), iii) training, iv) prompts and ChatGPT output as well as v) training ChatGPT with prompts.

3.1.1 Collection of training data, pre-processing of the data and training

15. The first three stages carry peculiar risks for the fundamental rights and freedoms of natural persons as “web scraping” enables the automated collection and extraction of certain information from different publicly available sources on the Internet (such as websites), which are then used for training purposes of ChatGPT.¹⁵ Such information can contain personal data which encompasses various aspects of the personal life of the respective data subject. Depending on the source, the scraped data may even contain special categories of personal data within the meaning of Article 9(1) GDPR.
16. Regarding web scraping, OpenAI brought forward Article 6(1)(f) GDPR as legal basis.¹⁶ It has to be recalled that the legal assessment of Article 6(1)(f) GDPR should be based on the following criteria:¹⁷ i) existence of a legitimate interest, ii) necessity of processing, as the personal data should be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed and iii) balancing of interests. Fundamental rights and freedoms of data subjects on one hand and the controller’s legitimate interests on the other hand have to be evaluated and balanced carefully.¹⁸ The reasonable expectations of data subjects should be taken into account in this assessment.¹⁹
17. As already stated by the former Article 29 Working Party, adequate safeguards play a special role in reducing undue impact on data subjects and can therefore change the balancing test in favor of the

¹¹ As described in para 8 of this Report.

¹² Judgement of the Court of Justice of 21 December 2023, C-667/21 (*Medizinischer Dienst*), para 79.

¹³ Regarding different stages of processing of personal data see Judgement of the Court of Justice of 29 July 2019, C-40/17 (*Fashion ID*), para 70.

¹⁴ For the purpose of this Report, the word “web scraping” covers both technical definitions of web scraping and web crawling.

¹⁵ For the purpose of this Report, it does not make a difference whether the entity scrapes personal data itself or acquires “scraped personal data” from a third party.

¹⁶ <https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-language-models-are-developed> (last accessed on 23 May 2024).

¹⁷ Judgement of the Court of Justice of 4 July 2023, C-252/21 (*Bundeskartellamt*), para 106.

¹⁸ EDPB Guidelines 3/2019 on processing of personal data through video devices V2.0, adopted 29 January 2020, para 17-40.

¹⁹ Recital 47 GDPR.

controller.²⁰ While the assessment of the lawfulness is still subject to pending investigations, such safeguards could *inter alia* be technical measures, defining precise collection criteria and ensuring that certain data categories are not collected or that certain sources (such as public social media profiles) are excluded from data collection. Furthermore, measures should be in place to delete or anonymise personal data that has been collected via web scraping before the training stage.

18. Regarding the processing of special categories of personal data, one of the exceptions of Article 9(2) must be applicable in addition, for the processing to be lawful. In principle, one of these exceptions can be Article 9(2)(e) GDPR. However, the mere fact that personal data is publicly accessible does not imply that “the data subject has manifestly made such data public”. In order to rely on the exception laid down in Article 9(2)(e) GDPR, it is important to ascertain whether the data subject had intended, explicitly and by a clear affirmative action, to make the personal data in question accessible to the general public.²¹
19. In the present context, where large amounts of personal data are collected via web scraping, a case-by-case examination of each data set is hardly possible. However, the aforementioned safeguards can contribute to meeting the requirements of the GDPR. For example, those measures should involve filtering data categories falling under Article 9(1) GDPR. The filtering should apply to both, data collection (for example, selecting criteria for what data is collected) and immediately after data collection (deleting data). In line with Article 5(2) and Article 24 GDPR, the burden of proof for demonstrating the effectiveness of such measures lies with OpenAI as controller.²²

3.1.2 ChatGPT input, output and training

20. The next stages concern ChatGPT input (including “prompts”), output and training.
21. Prompts refer to the input of data subjects when interacting with LLMs such as ChatGPT as well as file uploads and the user feedback regarding the quality of the data output²³ (the responses) of ChatGPT. OpenAI qualifies this as “Content” and publicly states to use this information to train and improve the model. In this context, Article 6(1)(f) GDPR is brought forward as legal basis.²⁴ OpenAI provides the option to opt-out of the use of “Content” for training purposes.
22. Data subjects should, in any case, be clearly and demonstrably informed that such “Content” may be used for training purposes. This circumstance is a factor to be taken into account in the context of the balancing of interests according to Article 6(1)(f) GDPR.²⁵

3.2 Fairness

23. It has to be recalled that the principle of fairness pursuant to Article 5(1)(a) GDPR is an overarching principle which requires that personal data should not be processed in a way that is unjustifiably detrimental, unlawfully discriminatory, unexpected or misleading to the data subject.²⁶ A crucial aspect

²⁰ Article 29 Working Party, Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC, WP 217, 844/14/EN, pages 31 and 42.

²¹ Judgement of the Court of Justice of 4 July 2023, C-252/21 (*Bundeskartellamt*), para 77.

²² Judgement of the Court of Justice of 14 December 2023, C-340/21 (*VB*), para 55.

²³ Regarding data output and the principle of data accuracy, see para 29 to para 31 of this Report.

²⁴ “Europe privacy policy” of OpenAI, Version 15 December 2023, effective by 15 February 2024, para 2 and 8.

²⁵ The TF members refer to the considerations regarding the balancing of interests in para 16 of this Report.

²⁶ EDPB Guidelines 4/2019 on Article 25 Data Protection by Design and by Default Version 2.0, adopted 20 October 2020, para 69.

of fairness is that there should be no risk transfer, meaning that controllers should not transfer the risks of the enterprise to data subjects.²⁷

24. With regard to ChatGPT, this means that the responsibility for ensuring compliance with GDPR should not be transferred to data subjects, for example by placing a clause in the Terms and Conditions that data subjects are responsible for their chat inputs.
25. Rather, if ChatGPT is made available to the public, it should be assumed that individuals will sooner or later input personal data. If those inputs then become part of the data model and, for example, are shared with anyone asking a specific question, OpenAI remains responsible for complying with the GDPR and should not argue that the input of certain personal data was prohibited in first place.
26. OpenAI already presented measures they put in place to address these issues.²⁸ However, it has to be recalled that the examination of those measures is subject to pending investigations at this point of time.

3.3 Transparency and information obligations

27. When web scraping personal data from publicly accessible sources such as websites, the requirements of Article 14 GDPR apply. Considering large amounts of data is collected via web scraping, it is usually not practicable or possible to inform each data subject about the circumstances. Therefore, the exemption pursuant Article 14(5)(b) GDPR could apply, as long as all requirements of this provision are fully met.²⁹
28. Contrary to this, when personal data is collected while directly interacting with ChatGPT, the requirements of Article 13 GDPR apply. In this context, it is of particular importance to inform data subjects that the aforementioned “Content” (the user input) may be used for training purposes.

3.4 Data Accuracy

29. In relation to the principle of data accuracy pursuant to Article 5(1)(d) GDPR, a difference should be made between input and output data. Input data can encompass either data collected, for instance, through web scraping or the “Content” provided by data subjects when using ChatGPT (such as “prompts”).³⁰ Output data encompasses the output following the interactions with ChatGPT.
30. It has to be noted that the purpose of the data processing is to train ChatGPT and not necessarily to provide factually accurate information. As a matter of fact, due to the probabilistic nature of the system, the current training approach leads to a model which may also produce biased or made up outputs. In addition, the outputs provided by ChatGPT are likely to be taken as factually accurate by end users, including information relating to individuals, regardless of their actual accuracy. In any case, the principle of data accuracy must be complied with.³¹

²⁷ EDPB Guidelines 4/2019 on Article 25 Data Protection by Design and by Default Version 2.0, adopted 20 October 2020, para 70.

²⁸ GPT-4 technical report, Dec. 2023, Chapter 6, arXiv:2303.08774v4 and Ouyang et al: “Training language models to follow instruction with human feedback”, Mar. 2022, arXiv:2203.02155v1.

²⁹ For further guidance regarding the requirements of Article 14(5)(b) GDPR see Article 29 Working Party, Guidelines on transparency under Regulation 2016/679, WP260 rev.01, endorsed by the EDPB.

³⁰ Regarding “Content” see para 21 of this Report.

³¹ Judgement of the Court of Justice of 16 January 2019, C-496/17 (*Deutsche Post AG*), para 57, according to which all processing of personal data must comply with the principles relating to data quality set out in Article 5 GDPR.

31. In line with the principle of transparency pursuant to Article 5(1)(a) GDPR, it is of importance that proper information on the probabilistic output creation mechanisms and on their limited level of reliability is provided by the controller, including explicit reference to the fact that the generated text, although syntactically correct, may be biased or made up. Although the measures taken in order to comply with the transparency principle are beneficial to avoid misinterpretation of the output of ChatGPT, they are not sufficient to comply with the data accuracy principle, as recalled above.

3.5 Rights of the data subject

32. The GDPR defines a set of rights of data subjects, for example to access personal data and being informed on how it is processed, to delete, to rectify, or under certain conditions to transmit personal data to a third party, to restrict the processing of the data subject's data or to file a complaint to an SA.
33. OpenAI as controller provides information on how to exercise these rights in its privacy policy (European version).³² In light of the complex processing situation and the factual limits for data subjects to intervene, it is imperative that data subjects can exercise their rights in an easily accessible manner.³³
34. In this context, OpenAI presents the possibility for contact by email, while some rights of the data subject can be exercised through the account settings. In line with Article 12(2) and Recital 59 GDPR, the controller shall continue improving the modalities provided for facilitating the exercise of the aforementioned data subject's rights.³⁴ In particular, this relates to the fact that, at least for the time being, OpenAI suggests users to shift from rectification to erasure when rectification is not feasible due to the technical complexity of ChatGPT.³⁵
35. As already mentioned above,³⁶ in line with Article 25(1) GDPR, the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate measures designed to implement data protection principles in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of the GDPR and protect the rights of data subjects.³⁷

4 ANNEX (QUESTIONNAIRE)

36. The following set of questions was developed within the context of the ChatGPT TF and is made available to the public. It has to be noted that SAs are independent and as such, each SA was free to modify the questionnaire or to add further questions. Moreover, differences in the questions may arise due to the respective official language to be used.

I. General

³² "Europe privacy policy" of OpenAI, Version 15 December 2023, effective by 15 February 2024, para 6 as well as the "Privacy policy" of OpenAI, Version June 2023, para 4.

³³ EDPB Guidelines 01/2022 on data subject rights - Right of access V2.0, adopted 28 March 2023, para 139-142.

³⁴ It has to be recalled that restrictions of these data subject rights are only permissible under the conditions specified in the GDPR.

³⁵ "Europe privacy policy" of OpenAI, Version 15 December 2023, effective by 15 February 2024 para 6.

³⁶ Para 7 of this Report.

³⁷ For further information on this topic, see the EDPB Guidelines 4/2019 on Article 25 Data Protection by Design and by Default Version 2.0, adopted 20 October 2020.

- a) Please provide a general and short description regarding your ChatGPT software infrastructure.
- b) Please provide a contact point of OpenAI. This is to ensure that we can contact you directly (and not via [redacted]), when necessary.
- c) Please provide the contact details of your Data Protection Officer (hereinafter: "DPO").

If you have not designated a DPO, please explain why not. In this context, please elaborate why the conditions of Article 37(1) GDPR do not apply for the processing of personal data carried out with the ChatGPT software infrastructure.

- d) In line with Article 30(4) GDPR, please provide a copy of your record of processing activities. This record can be limited to the processing of personal in relation to the product subject to this data protection audit, precisely the ChatGPT software infrastructure.

In any case, based on your record of processing activities, it must be clear for the [insert] Data Protection Authority which categories of personal data are processed for which purposes in the context of the ChatGPT software. Furthermore, it must be clear if and which special categories of personal data pursuant to Article 9 GDPR as well as personal data relating to criminal convictions and offences pursuant to Article 10 GDPR are processed.

- e) In line with Article 33(5) GDPR, please provide a copy of your documentation of personal data breaches.

II. Principles relating to processing of personal data

- a) Please describe, in a general manner, how you ensure that the principles relating to processing of personal data pursuant to Article 5(1) GDPR are complied with.

For example: Have you implemented a data protection management system (DPMS), do you carry out regular internal and/or external audits and, if applicable, is your designated DPO involved in all data protection matters? Please provide supporting documents, where applicable.

In your answer, please differentiate between the processing of personal data relating to users (individuals signing up to "ChatGPT services") and the processing of personal data collected from third parties (for example, from the Internet, via web scraping, to feed and train the algorithm).

- b) In line with Article 5(1)(b) GDPR, please describe the different purposes for which you process personal data in the context of the ChatGPT software infrastructure.

In your answer, please be precise about which data categories you process for which purposes.

If a precise description of the purposes of the processing can already be found in your record of processing activities (see question I. a) of this letter), please describe where in the record this can be found.

- c) In line with Article 5(1)(c) GDPR, please describe how you ensure that you limit the processing of personal data to what is necessary for your purposes.

- d) In line with Article 5(1)(d) GDPR, please describe how do you deal with the question of the accuracy of the personal data used and generated in the context of the ChatGPT software infrastructure?

In particular, how do you measure the accuracy of the personal data used in your language model training, testing and validation processes?

In this context, please provide the documents in which you describe the data accuracy measurement and assessment process both for the models' training, testing and validation data and for the models output.

- e) In line with Article 5(1)(e) GDPR, please describe how long you store personal data. If applicable, please provide a copy of your retention policy.

In your answer, please differentiate between the different data categories relevant for the ChatGPT software infrastructure.

If the time limits for erasure of the different categories can already be found in your record of processing activities (see question I. a) of this letter), please describe where in the record this can be found.

- f) In line with Article 5(1)(f) and Article 32 GDPR, please describe which technical and organizational measures are implemented in order to ensure appropriate security of the personal data processed in the context of the ChatGPT software infrastructure?

III. Data Protection Impact Assessment ("DPIA") and risk management

- a) Have you carried out a DPIA pursuant to Article 35 GDPR regarding the processing of personal data related to users and third parties in the context of the ChatGPT software infrastructure?

If so, please provide a full copy of your DPIA, as well as information when the DPIA has been carried out.

If not, please elaborate why the conditions of Article 35(1) and (3) GDPR do not apply for the processing of personal data carried out with the ChatGPT software infrastructure.

- b) If applicable, was your designated DPO involved when carrying out your DPIA pursuant to Article 35(2) GDPR?

If so, please provide a proof (for example, a copy of the statement of your DPO regarding the DPIA).

If not, please elaborate why not.

- c) Were you able to eliminate or adequately mitigate the risks identified when carrying out your DPIA?

Please provide documents as proof (for example, the copy of your risk-analysis).

If a precise description of the measures taken to eliminate or adequately mitigate the identified risks is part of your DPIA, please describe where in the DPIA this can be found.

- d) In line with Article 24(1) last sentence GDPR, which deadlines have been established for the periodic review of your DPIA?

- e) What is the age limit for the use of the ChatGPT software infrastructure and how do you ensure that these services are not used by data subjects (users) below this age limit?

IV. Lawfulness of processing

- a) Please describe from which different sources personal data are collected and then used for the training, testing and validation of the ChatGPT software infrastructure.

Please provide this information for all different stages of training and regarding the input and output through the usage by individuals.

- b) In line with Article 6(1) GDPR, please describe your legal basis (and, if applicable, your exceptions pursuant to Article 9(2) GDPR) for the processing of personal data in the context of the ChatGPT software infrastructure.

In your answer, please be precise and provide the respective legal basis for each different processing operation, in particular regarding the processing of users' personal data as opposed to processing of non-users' (third parties) personal data collected.

Please also specify whether and to what extent personal data communicated by the users via their interactions with the chatbot are used to train the AI system or for any other purposes by OpenAI, and if so, which is the applicable legal basis.

- c) Where applicable, please provide detailed information on the following and differentiate between the respective processing operations:

i) If the legal basis is consent pursuant to Article 6(1) (and possibly exceptions pursuant to Article 9(2)(a)) GDPR, please explain how consent is obtained and how the conditions for consent, in particular regarding Article 7 GDPR, are complied with.

ii) If the legal basis is the necessity for the performance of a contract pursuant to Article 6(1)(b) GDPR, please explain which contract with which content was concluded between which parties and why the processing of personal data is necessary for the performance of such contract.

Furthermore, if the legal basis is the necessity for the performance of a contract pursuant to Article 6(1)(b) GDPR, please explain which exception of Article 9(2) GDPR is additionally relied upon when special categories of personal data are processed.

iii) If the legal basis is legitimate interests pursuant to Article 6(1)(f) GDPR, please provide detailed information regarding the balancing of interest that you have carried out, in particular why you have concluded that such interests are not overridden by the interests of the data subjects.

Furthermore, if the legal basis is legitimate interests pursuant to Article 6(1)(f) GDPR, please explain which basis of Article 9(2) GDPR is additionally relied upon when special categories of personal data are processed.

iv) In case none of the above-mentioned legal basis are relied upon, please explain why another legal basis is applicable.

- d) Please explain why data subjects (users) have to enter their telephone number in addition to their email address?

Furthermore, please explain on which legal basis and for what purposes is the telephone number used and how long will the telephone number be stored by OpenAI?

- e) In line with Article 6(4) GDPR, does processing for a purpose other than that for which the personal data have been collected (“further processing”) take place?

If so, please describe which data categories are concerned by this and provide a copy of your computability test pursuant to Article 6(4) GDPR.

V. Rights of the data subject and Transparency

- a) Please explain how and when the information required pursuant to Article 13 and Article 14 GDPR is provided to the data subjects.

In the light of the currently available privacy policy published on your website, please explain in particular how and when the information required pursuant to Article 14 is provided to non-users (for example, third parties’ whose data are collected to train the algorithm relied upon).

Please provide screenshots (for example, of the part of your website where the data subject is informed) as well as a copy of your up-to-date privacy policy.

- b) Please explain how you ensure compliance with the rights of data subjects pursuant to Article 15 to Article 22 GDPR.

Please provide supporting documents, where applicable (for example, the copy of an internal policy dedicated to the handling of data subject requests).

- c) With particular regard to Article 17 GDPR, please explain how compliance with the right to erasure (“right to be forgotten”) is ensured?
- d) With particular regard to Article 21 GDPR, how do you comply with the requests of users that object to the processing of their personal data based on legitimate interest?
- e) With particular regard to Article 22 GDPR, does automated individual decision-making, including profiling, take place?

If so, please explain how compliance with Article 22 GDPR is ensured.

- f) Do you provide information to the users of your ChatGPT software infrastructure regarding the capabilities and limitations of the model and regarding the processing of personal data through language model services? If so, how?

VI. Transfers of personal data to third countries or international organisations

- a) Please, provide a list of the data centers you use to host and provide the ChatGPT software infrastructure (for example, to process personal data related to those services).

If personal data storage or processing location depends on some criteria, please describe them.

- b) In line with Article 44 GDPR, do transfers of personal data from users of [insert country] to third countries (outside the European Union) take place?
- c) If so, which instruments are being relied on for such transfers (Article 45, Article 46 and/or Article 49 GDPR)?

Please explain in detail why you have chosen the respective instrument and how compliance with the requirement of Article 44 GDPR – namely that the level of protection of natural persons guaranteed by the GDPR is not undermined – is ensured.

Please provide documents as proof (for example, if applicable a copy of the concluded Standard Data Protection Clauses pursuant to Article 46(2)(c) GDPR). In this context, please explain if and how you follow the recommendations given in the EDPB's Recommendations 01/2020 on measures that supplement transfer tools to ensure compliance with the EU level of protection of personal data.

VII. Disclosure of personal data to other parties

- a) Is OpenAI the only controller pursuant to Article 4(7) GDPR for the processing of personal data in the context of the ChatGPT software infrastructure?

If so, how is it ensured that no other party (for example, another company) decides the purposes and means regarding the processing of personal data in the context of the ChatGPT software infrastructure?

If not, who are the other (joint) controllers pursuant to Article 26 GDPR? In this case, please provide a copy of your arrangement pursuant to Article 26 second sentence GDPR.

- b) For the processing of personal data in the context of the ChatGPT software infrastructure, do you have a processor pursuant to Article 28 GDPR?

If so, please provide a copy of your contract pursuant to Article 28(2) GDPR.

- c) To which third parties are the personal data processed in the context of the ChatGPT software infrastructure disclosed to (for example, a new controller who processes the personal data generated and provided by OpenAI for its own purposes) and on which legal basis pursuant to Article 6(1) and/or exception pursuant Article 9(2) GDPR?

- d) When answering the questions of Point VII, please also take into consideration the integration of the ChatGPT software infrastructure into other products, such as (but not limited to) search engines.