



GP Partners Gawronski, Biernatowski Sp.K.
al. Jana Pawła II 12, 00-124 Warsaw | +48 22 243 4953 | info@gppartners.pl

Warsaw, November 24, 2023.

Complainant:
Łukasz Olejnik

Represented by:
r. pr. Maciej Gawronski
GP Partners Gawronski, Biernatowski Sp.K.
al. Jana Pawła II 12, 00-124 Warsaw

**President of the Office for Personal Data
Protection**
Stawki 2 Street
00-193 Warsaw

Case sign: DS.523.4800.2023.PR.SPI.

LETTER to supplement the position of Mr. Łukasz Olejnik in the complaint dated August 29, 2023.

Acting on behalf of the complainant, Mr. Łukasz Olejnik, we present a further position on legal issues related to the protection of personal data processed by generative artificial intelligence systems, including the ChatGPT tool provided by OpenAI.

We address this letter with reference to the Office for Personal Data Protection's communication of 20.09.2023 ("*Technology must comply with GDPR*") regarding the processing of the complaint of Mr. Łukasz Olejnik. We enthusiastically welcome the DPA's communication as a signal of the Authority's recognition of the legal issues raised in the complaint as relevant and of fundamental, universal importance for the rights of data subjects in connection with the further development of artificial intelligence tools. This is exactly how the subject matter of the case is viewed by Mr. Łukasz Olejnik. The complainant is more interested in assessing OpenAI's operating model for improving the level of protection of personal data processed under ChatGPT in the future, rather than in finding past violations.

Key legal issues. The Authority, in a communication dated 20/09/2023, notes that *the development of new technologies must respect the rights of individuals under, among other things, the GDPR*, and the case concerns OpenAI's systemic approach to European data protection principles.

In connection with these insights from the Authority, Mr. Łukasz Olejnik would like to highlight the issues that, in his opinion, play a key role in assessing OpenAI's systemic approach to data processing compliance with applicable regulations. These issues are:

- 1) Implementation of the obligation to ensure the protection of personal data in the design phase (**data protection by design**),
- 2) Ensure **the** exercise of the **right to rectify data** (for data used to train language models),
- 3) Implementation of **information obligations** relating to data processing for training of language models,

All of the above issues are manifestations of the fundamental principle, from the perspective of compliance with the GDPR, of processing personal data **lawfully, fairly and in a transparent manner** for the data subject, i.e. the principle enshrined in Article 5(1)(a) of the GDPR.



Selected issues to be investigated. In order to be able to analyze the compliance of ChatGPT's processing of personal data, OpenAI should provide the Authority with information in the following areas in particular:

- Whether a personal data processing impact assessment (DPIA) has been conducted for the ChatGPT tool and whether the assessment addresses the specific risks associated with the tool, including the risk of generating inaccurate content (known as hallucinations),
- How OpenAI has implemented mechanisms to handle the rights of the individual, in particular the right to information (Articles 13 and 14 of the GDPR), the right of access (Article 15 of the GDPR), and the right to rectify data (Article 16 of the GDPR),
- whether OpenAI has analyzed how the right to rectify data can be exercised, whether it has attempted to develop a technical solution to exercise this right,
- Whether OpenAI has analyzed the possibility of introducing a solution, a module to correct incorrect data, mechanisms to filter generated content, mechanisms for *machine unlearning*,
Such a module has been proposed for the Bard generative artificial intelligence tool provided by Google¹. The tool enables additional functionality for "second-hand" verification of content generated by artificial intelligence. The additional module, at the user's request, signifies, based on information in Google's browser, to what extent the generated content can be relied upon as to its reliability.
- Whether, in the event of an assessment that it is not possible to rectify personal data, OpenAI has put in place a policy to inform data subjects when and why rectification is not possible,
- what procedures OpenAI has in place in the event it becomes aware (e.g., as a result of a report) that certain content generated by the ChatGPT tool involving personal information is false,
- how OpenAI ensured that the ChatGPT tool was developed and released in compliance with the principle of privacy by design, what documentation helps ensure the principle of accountability in this regard.

Evaluation of the above issues requires OpenAI to describe the model and software on which the ChatGPT tool is based and analyse them in the context of relevant data protection regulations.

If the Office were to determine that OpenAI had conducted an analysis of the exercise of the right to rectification in particular, it would be reasonable to obtain documentation from OpenAI of this analysis and take it into account when investigating the complaint.

It would also be expedient for the resolution of the case, given the importance of the issues, to meet and exchange correspondence between the Office and representatives of OpenAI, with the participation of Mr. Łukasz Olejnik as the complainant and his attorney. This would allow us to understand what data protection challenges are associated with the development of generative artificial intelligence tools and how these challenges can be addressed in pursuit of a high level of protection of data subjects' rights and freedoms.

We would also like to submit for the Office's consideration an approach to the Office of Competition and Consumer Protection to seek its position on the consumer dimension of the issues covered by the complaint and the risks that artificial intelligence tools generate for consumer rights. Consumer issues in generative artificial intelligence tools are also of interest to foreign consumer protection regulators².

¹ <https://support.google.com/bard/answer/14143489?hl=en&co=GENIE.Platform%3DAndroid> (accessed 13.11.2023).

² The UK's Competition and Markets Authority (CMA), which is the country's regulatory watchdog, developed and then presented 7 principles for the responsible development and use of artificial intelligence tools on September 18, 2023 (Proposed principles to guide competitive AI markets and protect consumers, <https://www.gov.uk/government/news/proposed-principles-to-guide-competitive-ai-markets-and-protect-consumers>, accessed 13.11.2023).



Latest positions of authorities. Below we indicate additional positions and statements of authorities that have appeared recently and relate to the compliance of data processing using artificial intelligence tools.

- German data protection regulators have long been interested in OpenAI and the ChatGPT tool provided by OpenAI. Recently, the **Hessian Commissioner for Data Protection and Freedom of Information** (HBDI) announced that further detailed **questions** have been addressed to OpenAI regarding the conditions for processing data using the ChatGPT tool³.

According to the authority's press release, another 79 questions include the legality of bulk data collection for training language models and the ability of data subjects to exercise their rights to rectify data, erase data or obtain information regarding processing. The authority indicates that it is questionable whether these rights can be lawfully exercised at all in the context of ChatGPT processing.

The Hessian Data Protection Commissioner signals that he is taking supervisory action against OpenAI in cooperation with other national supervisory authorities, as well as EU authorities.

- In the complaint, in paragraph 1.4. (page 5 of the complaint), we described the French CNIL's announcement of intensified activity in the area of data protection in connection with generative AI tools. According to the announcement, the authority published **seven instruction sheets** on October 16, 2023 on **developing artificial intelligence systems**⁴. The worksheets are intended to support the design of artificial intelligence systems in accordance with data protection regulations.

The CNIL worksheets show, among other things, that the following issues should be considered when designing AI systems:

- (i) The need to conduct a **processing impact assessment** (DPIA). The assessment should cover, among other things, issues of risks to data subjects' rights, measures to allow the exercise of data subjects' rights, the level of transparency of processing.

The CNIL points out that one of the significant risks associated with AI systems is the risk of creating false content about a real person, which is particularly important in the case of generative AI systems and can have reputational consequences.

DPIA-based artificial intelligence system designers should decide on the need to implement specific technical measures such as (i) data minimization and (ii) mechanisms to facilitate the exercise of rights or other legal remedies, such as *machine unlearning* techniques or measures to explain and track the performance of AI systems.

³ Questions were addressed earlier (in April this year), and OpenAI provided answers to those questions. The *Hessian* authority, after analyzing the responses, determined that there was a need to ask additional, detailed questions to thoroughly investigate the compliance of the processing of personal data within the ChatGPT tool (*Hessischer Datenschutzbeauftragter fordert erneut Antworten zu ChatGPT*, 24.10.2023, accessed 13.11.2023)

<https://datenschutz.hessen.de/presse/hessischer-datenschutzbeauftragter-fordert-erneut-antworten-zu-chatgpt>

⁴ CNIL, *AI how-to sheets*, 16.10.2023, accessed 13.11.2023 (<https://www.cnil.fr/en/ai-how-sheets>).

- (ii) The need to take data protection into account in the **design of the architecture of artificial intelligence systems** - the CNIL indicates that it is necessary to choose an architecture that provides the highest level of protection of the rights and freedoms of data subjects, so that the processing complies with the principle of data minimization⁵.

The sheets published by the CNIL reflect the importance the authority attaches to the process of ensuring compliance of artificial intelligence systems at the design stage. The authority's position underscores the fundamentality of the principle of *data protection by design (privacy by design)* against the background of data protection regulation.

The sheets are currently undergoing public consultation, with the final version expected to be published in early 2024.

- On October 20, 2023. The **Global Privacy Assembly (GPA)**⁶ adopted a resolution on generative artificial intelligence systems. The GPA points out in the resolution:
 - (i) generative artificial intelligence systems must be designed, developed and implemented based on the principle of data protection, and data protection should be written into the concept and design of the system. The GPA thus emphasizes the role played **by the principle of privacy by design** and the need for a data protection impact assessment,
 - (ii) an important element of data protection compliance is **ensuring the accuracy of personal data** used to train models. GPA points out that one of the risks associated with GAI systems is the generation of false content containing personal data (so-called hallucinations). To mitigate this risk, among others, according to GPA, additional technical measures should be implemented - such as the use of filters for input and output data.
 - (iii) In terms of **transparency**, the GPA indicates, among other things, that data subjects should be able to obtain information on how, when and why personal data is used in the process of training generative artificial intelligence systems. The principle of transparency means that it should be transparent to know from what source the collection of data used to train models comes.
 - (iv) The GPA identifies as one of the key compliance aspects of GAI systems the obligation to **ensure that data subjects can exercise their rights**, including through the implementation of appropriate technical and organizational measures. Data subjects should be provided with, among other things, the right to access personal data as well as the right to correct inaccurate personal data.
- U.S. President Joe Biden on October 30, 2023 issued an executive order on "secure and trustworthy artificial intelligence." In the order, he calls on Congress to pass data privacy legislation to protect US citizens, especially children. Among other things, the order notes the need to support the acceleration of the development and use of privacy techniques, develop guidelines to assess the effectiveness of privacy protections, or evaluate how entities collect and use information available on the market⁷.

Conclusions. The above positions confirm that the issues raised in the complaint are timely and relevant, and the areas in which Mr. Łukasz Olejnik identifies irregularities coincide with the areas in which concerns and recommendations are formulated by the supervisory authorities.

⁵ CNIL, Taking data protection into account in the system design choices, 16.10.2023, <https://www.cnil.fr/en/taking-data-protection-account-system-design-choices>, accessed 13.11.2023)

⁶ The Global Privacy Assembly is a worldwide association of more than 130 regulators in the area of data protection and privacy.

⁷ FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/>



GP Partners

The essence of the matter and Mr. Łukasz Olejnik's perspective are particularly aptly described in the following passage from the GPA resolution, in which the GPA calls on entities providing generative artificial intelligence tools to:

to recognize data protection and privacy as a fundamental human right, and to build responsible and trustworthy generative artificial intelligence technologies that protect data protection, privacy, human dignity and other fundamental rights and freedoms⁸.

Enclosed we provide a machine translation of this letter into English as an aid.

Maciej Gawronski
Attorney-at-law

Attachments:

1. Machine translation of the letter

⁸ Self-translation.