The Hamburg Commissioner for
Data Protection and Freedom of Information

13.11.2023

Checklist for the use of LLM-based chatbots

Generative AI provided via a chatbot offers the opportunity to create content quickly and easily. Well-known large language models (LLM) are ChatGPT, Luminous and Bard. In many organisations, these tools have become part of the daily work, but often without binding guidelines for their use. The fact that the language models are usually operated in a cloud harbours various data protection risks. Firstly, the protection of confidential data is jeopardised because many companies work with the same cloud-based LLM model, inputs are used for further training of the models and business secrets as well as personal data may be transferred to them as a result. On the other hand, there is a risk of personal data being processed in an unauthorised manner due to incorrect results, especially in the case of particularly sensitive data categories. This checklist serves as a guide for companies and public authorities on the data protection-compliant use of chatbots.

1. Specify compliance regulations

Formulate and document precise internal directives as to whether and under what conditions which tools can be used. Specific examples of authorised and prohibited usage scenarios help to clarify this.

If you do not specify internal regulations on whether and how generative AI may be used in day-to-day work, you will have to assume that employees and other members of the organisation will make unauthorised and uncontrolled use of the new tools. The employing organisation may be liable for these actions.

2. Involve your data protection officer

Always involve your internal data protection officer when creating internal directives or implementing a use case for the first time. Depending on the use case, you should conduct a data protection impact assessment. If necessary, it may also make sense to get the works council on board.

3. Provide an organisation-owned account

Make non-personal chatbot accounts available in your organisation. Employees should not create an account on their own using private data. This would lead to profiling of the respective employee. If AI use in a professional context is desired, also professional accounts should be provided. If possible, these accounts should not contain the names of individuals. If the e-mail address is requested, it is advisable to specify an e-mail address created for this purpose only.

In some cases, mobile phone numbers are also requested during registration. In this case, it is also advisable to use a work telephone for this purpose. You should not authorise the private use of these business accounts.

4. Secure authentication

Accounts for AI chatbots used for business purposes offer considerable potential for abuse. If attackers gain unauthorised access to the application interface, they may be able to view previous activities if the chat history is not deactivated. They can also use their own queries to obtain personal information and business secrets. For this reason, a special focus must be placed on authentication. Use strong passwords and integrate additional authentication factors.

5. No input of personal data

If the terms and conditions allow the AI provider to process data for its own purposes, you should not transmit any personal data to the AI. This applies to any information that contains conclusions about customers, business partners or other third parties, as well as data of the company's own employees. It will generally not be possible to find a legal basis for this. The person entering the data must also not be identifiable if no legal basis can be found for the processing of their data.

6. No output of personal data

Make sure that the results of the AI application do not contain any personal data. Even if the input command does not name a person, the AI may include previous inputs or information from the internet under certain circumstances. Input should therefore be limited to cases that do not relate to individuals.

Example of an unproblematic input: "Write an advertising text for product X."

Example of a problematic input: "Which people are likely to be interested in product X?"

7. Caution with personally identifiable data

Also avoid entries that could possibly be related to specific persons. It is not sufficient to remove names and addresses from the entry. It may also be possible to draw conclusions about authors

and data subjects from the context. This risk is particularly high for AI applications that are designed to create cross-references from unstructured data.

Example: "Draft a job reference in the satisfactory range for a salesman at car dealership X." The entry can be personal if it is recognisable from which company it was made and at what time.

8. Opt-out of AI training

Reject the use of your data for training purposes. The manufacturers of AI models often use all the data entered for further training of their AI model. Private individuals and employees of other companies can then "ask" the AI to provide this content. However, depending on the service used, it is possible to object to the use for training purposes. In some cases, a specific contract model must be booked for this, which differs from the free standard application.

Example: With ChatGPT, for example, the opt-out is currently possible via the settings under

••• → Setting → Data Controls → Chat history and training

9. Opt-out of the history

Chat-based services often offer the opportunity to save previous entries so that the dialogue on a topic can be resumed at a later point in time. This inevitably means that a person's entries are linked to each other. The history should be disabled, especially if the dialogue is shared by several employees, as otherwise the content can be viewed by all colleagues. For the settings in ChatGPT, for example, see No. 8.

10. Check results for accuracy

The results of a chatbot request should be treated with caution. Large language models generate texts that come close to the desired result with mathematical probability. This does not mean that all the information provided is correct. On the contrary: the well-known LLMs usually base their results upon relatively old information. They are also known for the phenomenon of "hallucination", in which the AI invents statements that appear to be correct and logical but are actually incorrect. It is your responsibility as a user to check the accuracy of the result.

11. Check results for discrimination

Irrespective of their factual accuracy, results may also be inappropriate if, for example, they have a discriminatory effect. Data processing based on this may therefore be inadmissible because it violates anti-discrimination law or does not stand up to the balancing of interests under Art. 6 (1) (f) GDPR. You as the user are responsible for checking whether the answers are acceptable for further use within the legal framework.

Example: Information can also be discriminatory without any direct personal reference. An AI could make the following recommendation without naming individuals: "Male spectacle wearers should be favoured for the vacancy." Such a result could be based on an unauthorised analysis of health and gender data.

12. No automated final decision

Decisions with legal effect should generally only be made by humans. Otherwise, the requirements of Art. 22 GDPR must be met. If an LLM-based chatbot develops proposals that are accepted by employees, they must ensure that they have an actual scope for decision-making. Avoid being de facto bound to the suggestions due to the lack of transparency of the AI-supported preliminary work because you cannot understand the decision-making process. Insufficient resources and time pressure can also lead to results being accepted without being scrutinised.

13. Sensitise employees

Rise employees‘ awareness through training, guidelines and discussions as to whether and how they are allowed to use AI tools.

14. Data protection is not everything

The protection of personal data must not be undermined by the use of AI services. It is also advisable to regulate other aspects such as the protection of copyrights or trade secrets. In the case of official use cases, disclosure bans and other regulations must be considered.

15. Follow further developments

The regulation of artificial intelligence is currently being prepared at EU level. The future AI regulation is likely to affect not only the providers of such services, but also certain users. Due to advancing technical solutions and ongoing updates to new systems and language models, it should be regularly reviewed whether the internal requirements need to be adapted.

In addition, the data protection authorities are currently conducting precedent setting procedures with the aim to find out if the LLM on the market are lawful in general.