Guidance on AI and Data Protection

1. Introduction

AI applications are an exciting technology used by millions of users worldwide. To continuously improve these self-learning applications/platforms, a constant learning process typically occurs, including the data provided by users. Often, a significant amount of personal data provided by users is processed within these systems. Therefore, caution is advised from a data protection perspective when entering personal, sensitive, or confidential company data.

The following guidance provides an overview of AI, measures, and specific application areas for using AI:

2. General Definition of AI (Artificial Intelligence)

In the regulation of the European Parliament and Council on artificial intelligence, AI is defined as: "Artificial intelligence system" (AI system) means software designed to achieve a set of goals defined by humans, producing outputs such as content, predictions, recommendations, or decisions that influence the environment with which it interacts."

3. Application Areas

AI applications can be used in various areas at LUNEX, such as:

- Chatbots
- Assistance in research and teaching
- Translations
- Writing and improving texts and programming code
- Idea generation
- Support in content creation
- Automation of customer interactions

4. Free / Paid AI Application Offers

AI applications often offer the same or similar functionalities in free and paid versions. In the free version, users often "pay" with the data they input, which the provider uses to improve the system. In contrast, in the paid version, the input data remains in the system only for the duration of use and is immediately deleted without further use by the provider. Therefore, it is essential to check the provider's data usage policy before using an AI application.

5. General Handling of Personal Data in AI Applications

AI applications (such as ChatGPT) are typically self-learning systems. Improving these systems continuously requires extensive training.

- This involves feeding in or having the system process vast amounts of data.
- Furthermore, depending on the application (e.g., free or paid versions), user input can often be used to "train" the AI. In this process, a significant amount of personal data provided by users — such as uploaded documents — may be processed by the systems.

From a data protection perspective, it is essential to first check how the input data will be further processed for the provider's own purposes, such as training the AI.

Below, the general data protection approach for the use of AI applications is presented.

a) Anonymization

If the AI application uses data for its purposes, such as "training the AI," anonymization of the data is required from a data protection perspective. All personal data (e.g., name, student ID, address) and identifiable data (e.g., specific work phrasing, membership in small workgroups/test groups) must be removed from the input and document to ensure no personal reference is possible. Additionally, confidential company data should not be provided.

b) Data Minimization

Personal data processed in an AI application should be limited to the necessary amount. The principle of data minimization should be applied.

c) Purpose Limitation

Under data protection laws (GDPR, BDSG), there is a purpose limitation for processing personal data. Personal data may only be processed within the scope of the intended purpose, e.g., through consent or contractual obligations (such as a student contract) in an AI application.

d) Use/Evaluation of AI-Generated Results

Due to insufficient data foundations and designs, AI results may negatively impact and violate the rights and freedoms of affected individuals. Additionally, these results may be incorrect or violate GDPR requirements, such as the principles of fairness, data accuracy, legitimate purposes, or proportionality. Before using AI results, users should evaluate their accuracy and appropriateness, ensuring they do not violate the rights and freedoms of affected individuals.

e) Automated Decision-Making

Data protection laws specifically regulate "automated decision-making." It must be ensured that AI results do not have direct and immediate effects on affected individuals but are reviewed for accuracy and applicability by natural persons.

f) Responsibility

When using an AI application, users must take responsibility for the application and implement necessary measures to ensure lawful processing, protection of affected individuals' rights, processing security, and control of the AI application.

6. Specific AI Applications at LUNEX

a) Guidance for Lecturers

A guidance document on AI and teaching has been created and published for staff and students.

b) Translation Tool "DeepL"

Paid Version: The paid DeepL Pro service guarantees that DeepL never stores the texts translated by DeepL Pro subscribers. Personal data can be processed.

Free Version: In the free version, all personal and sensitive company data must be removed for data protection reasons

c) ChatGPT Application

ChatGPT uses input data for training the AI application without an opt-out option. Therefore, it is advised not to process personal or sensitive company data in the system for data protection reasons.

Additionally, ChatGPT transfers all liability and legal compliance to the user. For example, uploading copyrighted material to ChatGPT requires necessary permissions. If ChatGPT is no longer used or employees leave the company, the account and corresponding data should be deleted.

Data protection requests can be submitted at the following link: OpenAI Privacy Center.

Furthermore, data input history can be requested or exported at the following link: How do I export my ChatGPT history and data? | OpenAI Help Center.

d) Microsoft 365 Copilot Chat

The paid Bing Chat Enterprise application ensures that session content is not used for training the AI. Microsoft explains this in detail in Microsoft 365 Copilot Chat Privacy and Protection | Microsoft Learn.

Due to the assurances provided in Protecting Company Data in Microsoft 365 Copilot and Microsoft 365 Copilot Chat | Microsoft Learn, it is, from a data protection perspective, permissible for users to use Copilot within the scope of enterprise licenses with personal data in accordance with data classification guidelines.

It should be noted, however, that highly sensitive personal data such as health data (according to the data classification) should not be processed in the application.

This guidance may be gradually expanded within the scope of the application. If you are using any additional applications, please feel free to contact the data protection team at: cnadimi@lunex.lu.