



**CAMPUS
TAGE** 2024

KI trifft auf Datenschutz Digitalstrategien




DSN
Aka
demie

Privacy in ChatGPT and Copilot – A deep dive into some of the most popular AI tools

Tania Vanessa Eslava Suárez
www.first-privacy.com

Agenda

- 1 What is AI?
- 2 Privacy in the context of ChatGPT and Copilot



1. What is AI?

Artificial Intelligence

"science and engineering of making intelligent machines, especially intelligent computer programs"

McCarthy, What Is Artificial Intelligence? (2007, p.2)

AI is not a new topic

In 1946, **Alan Turing** suggested that a machine could simulate mathematical deduction and formal reasoning.

It was founded as an **academic discipline** in 1956

In 1997, **Deep Blue** won a chess match against Garry Kasparov.

AI has been integrated in computer science ever since

Were calculators and computers **artificial intelligences** all along?

No! Under our current understanding of AI, we need **Machine Learning**

Machine Learning

"[...] computer systems that are able to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyse and draw inferences from patterns in data."

Definition from Oxford Languages

Traditional Systems are rule-based

They work by following a set of pre-determined instructions.

Instruction:

If search = „Bogotá“
Show results that contain the word „Bogotá“.

Machine learning works with statistics

Large volumes of training data
+
Computing power
=
How probable is something

The training will show that „Bogotá“ appears in the context of: a city, the country of Colombia, Latin America, and the Netflix Show „Money Heist“ which has a character named „Bogotá“.

Depending on the query, the AI will be able to identify the most appropriate answer.

- „Where is Bogotá“: Most probable answer lays in the context of Colombia, Latin America, etc.
- „Who plays Bogotá“: Most probable answer lays in the context of the Netflix show.

Large Language Models

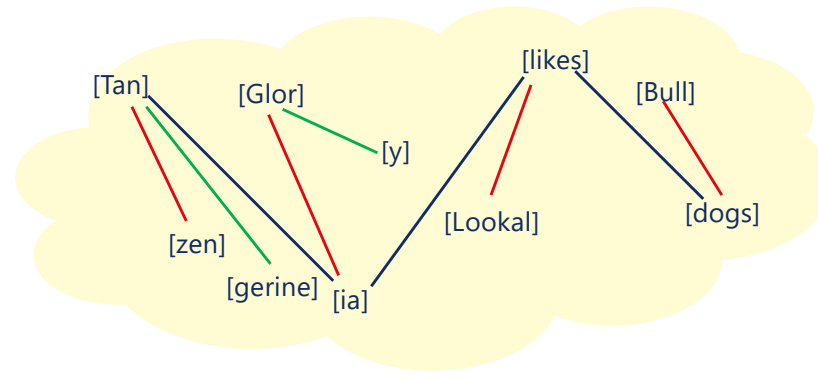
Specifically designed to understand and generate human-like text


Tania likes dogs

[Tan] [ia] [likes] [dogs]

Large Language Models


Specifically designed to understand and generate human-like text





Tell me something about Tania

[t] [ri] [athlon]
[anz] [en]
[oma] [to]
[ania] [mili] [meter]
[likes] [dogs]
[Ray] [mond] [quark]
[pans]



Tell me something about Tania

Tania likes dogs!

Difficulty of the German language: "Ich glaube, dass sie jetzt den grünen Apfel isst."



2.

Privacy in the Context of ChatGPT and Copilot

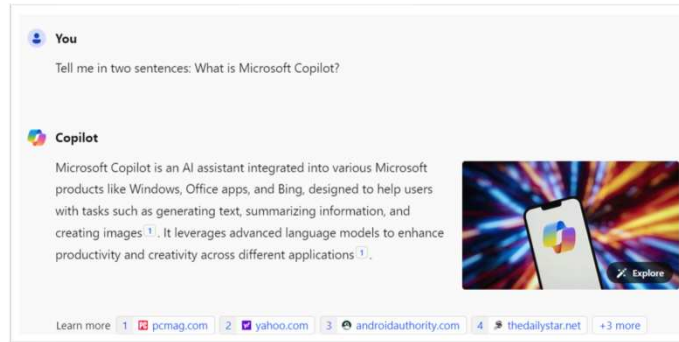


What is ChatGPT?

Tell me in two sentences: what is ChatGPT

ChatGPT is an AI language model developed by OpenAI, designed to generate human-like text based on the input it receives. It excels in a variety of tasks, including answering questions, writing creative content, and engaging in detailed conversations across numerous topics.

What is Copilot?



Training

Performed by the company that developed the AI model („provider“).

Use

Performed by the company that contracted the service („deployer“) through their employees.

Training

Performed by the company that developed the AI model („provider“).

Option 1: train without personal data.

Option 2: comply with all the obligations and requirements for a lawful processing of personal data. Particularly important here: **transparency** and having a **legal basis**.

Potential violations during the LLMs training do not affect the lawfulness of using such a model within an AI system

Training

Performed by the company that developed the AI model („provider“).

Use

Performed by the company that contracted the service („deployer“) through their employees.

FIRST PRIVACY

Privacy problems emerge with AI's inputs and outputs

FIRST PRIVACY

Input is the data or information provided to an AI model or system.

Output is the result produced by the AI system after processing the input data.

FIRST PRIVACY

ChatGPT

😊 *Please give me ideas for a fruit company slogan*


Sure! what about "Nature you can trust". 🌀

FIRST PRIVACY

ChatGPT

Please draft a formal message to express support to my client, Mr. Hosokawa, regarding his recent cancer diagnosis.

ChatGPT – Personal Accounts




Careful! When using personal accounts, information is at risk.

- OpenAI may use the data to improve their models.
- Any data contained in the conversations can potentially be read by OpenAI employees.
- Conversations cannot be deleted from their backend.

Personal data M&A Plans Financial information

Trade secrets Intellectual Property

ChatGPT Team/Enterprise




Business versions provide enhanced **security**, a **Data Processing Addendum** (+SCCs), insurance policies, and more.

Please note that the DPA is not automatically executed upon service acquisition; **customers must specifically request it.**

Who can read conversations in ChatGPT business accounts?

- Within the customer's organization, only **end users.**
- OpenAI employees to resolve incidents, recover end user conversations (with permission), or where required by law.

ChatGPT Team/Enterprise



Please draft a formal message to express support to my client, Mr. Hosokawa, regarding his recent cancer diagnosis.

📧 Certainly. Here's a draft for your formal message:

Subject: Offering My Support During This Challenging Time

Dear Mr. Hosokawa,

I was deeply saddened to hear about your recent diagnosis. Please know that you have my full support during this challenging time.


If there is anything I can do to assist you, whether professionally or personally, please do not hesitate to

Sharing accounts?


ChatGPT 4o

- Model
- GPT-4o For daily tasks
- GPT-4o mini Fastest
- GPT-4 Legacy model
- Temporary chat

What about Outputs?



Outputs in ChatGPT depend mostly on the training data. But outputs on Copilot can quickly become a privacy problem.




Copilot is designed to access a broad spectrum of data within the Microsoft 365 ecosystem

- Emails
- Documents
- Meetings
- Chats
- Calendars
- Contacts






Copilot operates under a company's established data governance frameworks.


*There is a risk of **information overreach**: Employees might inadvertently receive access to sensitive or irrelevant information due to the extensive data integration of Copilot*



AI's privacy problems are variations of old problems





Copilot highlights Microsoft's extensive access to data, drawing attention to potential issues arising from misuse or negligent handling of Microsoft 365 features.



Mitigating Risks through Data Governance

- Detailed access controls and regular audits
- Data retention policies and management of the lifecycle of information
- Use of sensitivity labels



Conclusions:

- Traditional systems follow instructions, **machine learning** identifies **patterns** and relationships in data.
- When using AI models, always think of what can go wrong with the **input** and with the **output**.
- Privacy problems with AI are just variations of old problems. Good data protection practices go a long way in addressing them.



Vielen Dank für Ihre Aufmerksamkeit!

Konsul-Smidt-Straße 88 • 28217 Bremen
T +49 (0) 421 69 66 32-298
Mail: akademie@dsn-group.de

<https://www.dsn-group.de/dsn-akademie>