

Information Security and Artificial Intelligence (AI)

The use of AI systems such as ChatGPT or Copilot offers opportunities, but also involves risks. On the one hand, productivity can be increased and problem-solving can become easier. On the other hand, there is a risk of data leaks, privacy violations, or manipulation of AI-generated results. The goal when using AI should therefore be to maximise the benefits while minimising the risks. You can find the security guidelines for the use of AI services in the following flyer.

It is the responsibility of the users to use external AI services in a purposeful and responsible manner. It is not permitted to share confidential or strictly confidential data on these platforms. This includes the prohibition of using particularly sensitive or otherwise personal data.

WHAT IS generative AI?

Artificial intelligence (AI) refers to the ability of machines to perform tasks that typically require human intelligence, such as learning, problem-solving, and decision-making. Generative AI can create new content, such as text, images, or music. These systems use machine learning and neural networks and are trained on large amounts of data. They can also be specialized for specific tasks, such as translations.

Responsible Use of AI

Use AI in practice to learn how best to apply the technology for your work. Experiment with different AI tools and use AI to automate routine tasks. Here are some things you should and should not do:

Do's



- Write meaningful prompts
- Critically review the AI results
- Turn off the history function (if possible)
- Absolutely consider copyright when using AI-generated content

Don'ts



- Don't share confidential information with AI systems
- Don't share personal data
- Don't blindly rely on AI results
- Don't ignore the ETH guidelines and data protection regulations

Additional information on Do's:

1 Write meaningful prompts

When creating prompts for conversational AI like ChatGPT, follow these recommendations to ensure the best possible results and the security of your and ETH information:

1. Be clear, provide context; use keywords; pay attention to correct spelling; define the AI's role and the desired output.
2. Do not use personal, confidential or proprietary information.
3. Use respectful language.
4. Promote diversity and inclusion.
5. Regularly evaluate and refine your prompts.

More information on prompting:
<https://ethz.sharepoint.com/sites/ArtificialIntelligence>

2 Critically review the AI results

Read and evaluate the output of generative AI to assess the quality, accuracy, and ethical implications of the generated content before reusing it.

Caution:
AI can sometimes generate false information, the so-called "hallucinations". This happens when the model creates content for which it has no training data. These false information can appear very convincing. Therefore, you should always verify the links and sources provided by the AI.

3 Turn off the history function (if possible)

Certain AI solutions offer options to restrict the storage and use of the provided data. There may be various control and adjustment levels, use them to prevent your data from being used uncontrollably on the web.

4 Absolutely consider copyright when using AI-generated content

The ability of AI to generate original content raises complex copyright issues. ETH has developed clear guidelines for the use and dissemination of AI-generated content, especially in the area of teaching and learning.

You can find them here:
<https://ethz.ch/en/die-eth-zuerich/lehre/ai-in-education.html>

ETH has introduced a system for classifying information. It distinguishes between public, internal, confidential, and highly confidential data. AI systems must be developed accordingly to this classification and the data must be used appropriately to ensure the right level of protection.

STRICTLY CONFIDENTIAL
CONFIDENTIAL
INTERNAL (DEFAULT)
PUBLIC

Additional information on Don'ts

1 Don't share (strictly) confidential information with AI systems

In addition to the benefits of artificial intelligence, please also consider its potential risks

The incorrect use of AI can pose several dangers for ETH:

1. Reputational damage
2. Ethical concerns
3. Financial losses

Examples of confidential and strictly confidential data:

- Research data & research projects
- Strategic & financial documents
- Contracts & agreements
- Confidential information & protected intellectual property
- IT & security-sensitive data
- Crisis management & emergency documents
- Corporate & organizational projects

2 Don't share personal data

"Personal data" means any information relating to an identified or identifiable natural person ("data subject").

An identifiable natural person is someone who can be directly or indirectly recognized. This can occur through information such as a name, identification number, location data, or specific characteristics that reveal physical, genetic, psychological, economic, cultural, or social attributes of that person.

"Particularly sensitive personal data" are information about:

- Data concerning religious, ideological, political, or trade union views or activities
- Data concerning health, intimate life, or racial or ethnic origin
- Genetic data
- Biometric data that uniquely identifies a natural person
- Data concerning administrative or criminal prosecutions or sanctions
- Data concerning measures of social assistance

3 Don't blindly rely on AI results

Be thorough in checking AI-generated texts.

When you've asked AI to create content, watch out for:

1. Lack of coherence:
AI-generated texts can lack logical flow, leading to confusing passages.
2. Semantic errors:
AI models can occasionally misinterpret context, resulting in errors, inaccuracies, or implausible statements.
3. Repetitions:
AI-generated content can exhibit repeated patterns or redundant information that impair the quality and uniqueness of the text.

4 Don't ignore the ethical guidelines and data protection regulations

AI systems at ETH must be developed and operated in accordance with the DSG, which includes principles such as transparency, purpose limitation, data minimization, and the right to information and deletion.

ETH has developed a specific strategy for the use of AI in teaching that takes into account the following aspects:

1. Promoting critical thinking
2. Ethical use
3. Adaptation of teaching methods
4. Preparation for the job market

For questions about information security

Contact the Information Security Officer (ISO) responsible for your department or the CISO.

Contact persons list:
<https://ethz.ch/staffnet/en/service/information-security/contacts.html>