



## AI and research data

The use of AI in qualitative research seems very convenient, especially when processing and evaluating the data collected. Some transcription services use AI, software for computer-assisted qualitative data analysis (e.g. MAXQDA, ATLAS.ti) often have AI-supported functions, other tools such as ChatGPT offer data analysis directly.

### What should I consider when using AI on research data?

**Data protection** always applies when so-called personal data are processed. This includes any data relating to a person who can be directly or indirectly identified based on these data. The processing of personal data is generally **prohibited**.

There are two **exceptions**:

1. If the affected person **consents to data processing**: If you, for example, want to use AI tools for transcribing an interview, you have to obtain the interviewee's consent prior to data collection as part of the **declaration of informed consent**.
2. If a specific **legal requirement** applies: The University of Vienna, for example, uses the learning platform Moodle to meet its statutory responsibilities. Therefore, students do not have to consent to the processing of their data when accessing different activities on Moodle. This is covered by the Data Protection Declaration of the University of Vienna.

**What should I consider when using AI tools for data processing and analysis?** If you want to use AI tools to process your data (e.g. to prepare interview transcripts) or to analyse data, consider the following points:

- **Never** enter your own or another person's personal data or confidential data in an AI tool.
- Make sure that it is not possible to identify any person(s) from the context: You can also identify persons based on contextual information only. This is why you should prepare a plan for **pseudonymisation**.
- **Inform** research participants about the use of AI tools in the **declaration of informed consent** and obtain explicit **consent**.
- Explain what **providers of AI tools** do with the data entered and where they store the data (especially when using free tools): If the providers do not process the data in Europe, data protection according to EU standards cannot be guaranteed. If you, for example, provide an US-based AI company with a text, you often share your (own or another person's) data without permission.
- If you use an AI tool without entering personal data, you still have to maintain **confidentiality**. You should generally not enter confidential data, company secrets or unpublished research results in an AI tool.
- An exception to this is the **local** use of an AI-supported tool (i.e. on your computer) or if the terms of use **guarantee** that the data entered are **not stored** and are not used for training purposes.
- **The requirements of the course apply**: The teacher decides on (non-)permitted materials in the course. Everything that is not explicitly permitted is forbidden.



### What should I consider and which decisions do I have to take?

- **Limited control over the analysis process:** The results of AI-supported data analysis are not always transparent.  
Generative AI tools are language models with no consciousness. They are based on linguistic data and thus simulate human language. This is why AI may generate biased, unclear and incorrect information that seems plausible. Always also check the AI output by randomly analysing your data yourself.
- **Document your research process** (e.g. in a [research journal](#), PDF in German) to make your decisions transparent.
- **Acquire a new competence:** The ability to understand and to use AI is a competence that you can learn. Therefore, you have to engage in continuous learning to ensure the best possible use of this technology.
- **You are responsible for your decisions:** You have to decide whether you want to use generative tools or not.  
Ultimately, you are responsible for your learning process.
- **You remain the author of your text:** You can also decide on **not** using AI.