



KI und Forschungsdaten

Im qualitativen Forschungskontext scheint der Einsatz von KI gerade bei der Aufbereitung und Auswertung von erhobenen Daten sehr praktisch. Einige Transkriptions-Services nutzen KI, Software zur computergestützten qualitativen Datenanalyse (z.B. MAXQDA, Atlas.ti) hat häufig bereits KI-gestützte Funktionen, andere Tools wie ChatGPT bieten die Auswertung von Daten an.

Was ist beim Einsatz von KI bei Forschungsdaten zu beachten?

Datenschutz greift immer dann, wenn sogenannte personenbezogene Daten verarbeitet werden. Damit sind alle Informationen gemeint, die sich auf eine Person beziehen und diese dadurch direkt oder indirekt identifiziert werden kann. Die Verarbeitung von personenbezogenen Daten ist grundsätzlich **verboten**.

Es gibt zwei **Ausnahmen**:

1. wenn die betroffene Person **einwilligt**: wenn Sie zum Beispiel KI-Tools zum Transkribieren eines Interviews verwenden möchten, müssen Sie die Zustimmung vor der Datenerhebung im Rahmen Ihrer **informierten Einwilligungserklärung** einholen.
2. wenn eine bestimmte **gesetzliche Anforderung** vorliegt: die Verwendung der Lernplattform Moodle ist für die Universität zum Beispiel notwendig, um ihre gesetzlichen Aufgaben zu erfüllen. Deshalb müssen Studierende nicht extra zustimmen, dass ihre Daten dort verarbeitet werden. Das ist über die Datenschutzerklärung der Universität Wien abgedeckt.

Worauf muss ich beim Einsatz von KI-Tools zum Aufbereiten und Auswerten von Daten achten?

Wenn Sie KI-Tools zum Aufbereiten Ihrer Daten (z.B. Erstellen von Interviewtranskripten) oder zur Datenanalyse verwenden wollen, achten Sie darauf:

- **nie** eigene oder fremde personenbezogene Daten oder vertrauliche Informationen in ein KI-Tool einzugeben.
- dass auch aus dem Kontext nicht auf die Person(en) geschlossen werden kann: Man kann Personen auch aufgrund von Kontextinformationen identifizieren. Legen Sie sich deshalb einen **Pseudonymisierungsplan** zurecht.
- dass Sie Teilnehmer*innen an Ihrer Forschung über die **informierte Einwilligungserklärung** über den Einsatz von KI-Tools **informieren** und sie diesem explizit **zustimmen**.
- was die **Anbieter*innen von KI-Tools** mit den eingegebenen Daten machen und wo diese gespeichert werden (insbesondere bei Gratis-Tools): Wenn Daten nicht in Europa verarbeitet werden, kann der Datenschutz nach EU-Standards nicht gewährleistet werden. Wenn man z.B. einen Text in eine KI-Anwendung mit Sitz in den USA hochlädt, gibt man häufig zugleich die (eigenen oder fremden) Daten ohne Erlaubnis weiter.
- Verwenden Sie ein KI-Tool, ohne personenbezogene Daten einzugeben, stellt sich immer noch die Frage der **Vertraulichkeit**. Vertrauliche Daten, Firmengeheimnisse, unveröffentlichte Forschungsergebnisse sollten grundsätzlich nicht in ein KI-Tool eingegeben werden.
- Eine Ausnahme bildet der Fall, wenn Sie ein KI-gestütztes Tool **lokal** verwenden, also nur auf Ihrem Rechner oder das KI-gestützte Tool in den Nutzungsbedingungen **garantiert**, dass die eingegebenen Daten **nicht gespeichert** und für Trainingszwecke weiterverwendet werden.
- **Die Vorgaben der Lehrveranstaltung gelten**: In der Lehrveranstaltung entscheidet die Lehrperson über erlaubte oder nicht erlaubte Hilfsmittel. Alles, was nicht explizit erlaubt ist, ist verboten.



Was muss ich bedenken und welche Entscheidungen muss ich treffen?

- **Begrenzte Kontrolle über den Auswertungsprozess:** Die Ergebnisse einer KI-gestützten Auswertung von Daten ist nicht immer nachvollziehbar. Generierende KI sind Sprachmodelle ohne Verständnis; sie werden auf Sprache trainiert und simulieren menschliches Sprechen. So kann KI verzerrte, unklare und fehlerhafte Informationen generieren, welche trotzdem plausibel klingen. Überprüfen Sie Ergebnisse immer, indem Sie Ihre Daten auch selbst mindestens stichprobenartig auswerten.
- **Dokumentieren Sie Ihren Forschungsprozess,** z.B. in einem [Forschungsjournal](#), um Ihre Entscheidungen nachvollziehbar zu machen.
- **Sie eignen sich eine neue Kompetenz an:** Die Fähigkeit, KI zu verstehen und zu nutzen, ist selbst eine Kompetenz, die erlernt werden muss. Daher ist ein kontinuierlicher Lernprozess erforderlich, um die bestmögliche Nutzung dieser Technologie zu gewährleisten.
- **Sie treffen die Entscheidungen:** Sie müssen die Entscheidung treffen, ob Sie generative Tools einsetzen wollen oder nicht. Letztendlich sind Sie für Ihr eigenes Lernen verantwortlich.
- **Sie bleiben Autor*in Ihres Textes:** Sie können sich auch dafür entscheiden, KI **nicht** zu nutzen.