

Anhang Deklarationspflicht von generativer KI bei Arbeiten

1. Anwendungsbereich

Dieser Anhang (nachfolgend «Anhang») zur [Richtlinie KI bei Leistungsnachweisen](#) (nachfolgend «RL KI») dient ausschliesslich dazu, die aus Ziff. 6.2 RL KI resultierende Deklarationspflicht für generative KI-Systeme (vgl. zur Definition Ziff. 5 RL KI) verbindlich festzulegen, sofern die Verantwortlichen (vgl. zur Definition Ziff. 3 RL KI) den Umfang und den Detaillierungsgrad der Deklarationspflicht gemäss den fachspezifischen und prüfungsdidaktischen Anforderungen nicht präzisieren.

Es handelt sich demnach um subsidiär¹ zur Anwendung gelangende Bestimmungen. Es liegt in der abschliessenden Verpflichtung der Verantwortlichen sicher zu stellen, dass die Studierenden bzw. Weiterbildungsteilnehmenden rechtzeitig und in geeigneter Form Klarheit darüber erhalten, ob Präzisierungen vorliegen oder ob in Abwesenheit eigener Anforderungen der vorliegende Anhang die Deklarationspflicht von generativen KI-Systemen bei Arbeiten regelt.

2. Abgrenzung

Der vorliegende Anhang befasst sich prioritär mit Text- und Bildgenerierung durch generative KI-Systeme und nicht mit Themen wie «KI und Programmiercode». Sofern die Verwendung von Code-Assistenten o.ä. in Arbeiten wesensrelevant ist, sind die Verantwortlichen in der Verpflichtung, die Deklarationspflicht näher zu umschreiben. Gleichermassen nicht erfasst werden Prüfungen (vgl. zur Definition Ziff. 6.1 RL KI). Die Verantwortlichen sind bei Prüfungen angehalten, die Verwendung von generativen KI-Systemen gleichermassen selbständig zu regeln und zu kommunizieren.

3. Deklarationspflicht

Generative KI-Systeme müssen wie andere Hilfsmittel und Quellen angegeben werden, wenn Output (vgl. zur Definition Ziff. 5 RL KI) in die Arbeit eingearbeitet wird. Der Anteil bzw. das Ausmass des Mitwirkens von generativen KI-Systemen an der schöpferischen Leistung einer Arbeit muss für Dritte erkennbar sein.

Im Anwendungsbereich dieses Anhangs gelten folgende Deklarationspflichten:

- I. Ein generatives KI-System ist **im Text der Arbeit zu zitieren**, wenn Inhalte (Text, Bilder, Daten oder andere), die mit diesem Tool erstellt wurden, paraphrasiert, zitiert oder in die eigene Arbeit sinngemäss eingebaut werden (vgl. bereits Ziff. 6.2.1 RL KI).

Zitiervorgabe²:

1. Autor des Modells (z.B. OpenAI als Autor des Large Language Models ChatGPT)
2. Datum (i.e. das Jahr in welchem ein generatives KI-System genutzt wurde (und nicht das exakte Verwendungsdatum))
3. Bildbezeichnung (bei durch generative KI-Systeme generierten Bildern ist der (Initial-)Prompt als Bildbezeichnung in Anführungs- und Schlusszeichen anzugeben)

¹ Die Bestimmungen des Anhangs gelten also «hilfsweise», wenn keine anderen festgelegt wurden.

² Quelle: <https://apastyle.apa.org/blog/how-to-cite-chatgpt>, abgerufen am 27.10.2023.

Beispiele (In-Text-Referenz):

- Wörtliche Übernahme von Output:
«Eine Namensänderung ist die rechtliche Veränderung des Namens einer Person» (OpenAI, 2023).
- Paraphrasierung von Output:
Die Namensänderung kann aus verschiedenen Gründen wie Heirat, Scheidung, Adoption, oder achtenswerten Gründen vorgenommen werden (OpenAI, 2023).
- Übersetzung:
«Un époux ne peut être autorisé à changer de nom selon l'art. 30 al. 1 CC que dans la mesure où le(s) nom(s) correspond(ent) aux règles sur le nom prévues par le droit de famille (art. 160 et art. 30 al. 2 CC)» (DeepL, 2023).
- Bild:



«Die unendliche Fülle der Namensänderung» (DALL-E 2, OpenAI, 2023).

- II. Werden generative KI-Systeme wie unter Ziff. 3.I verwendet oder anderweitig zur Texterstellung (z.B. Literaturrecherche oder Strukturierungshilfe) oder zur (Weiter-)Bearbeitung eigener Daten oder bereits generiertem Output genutzt, müssen die verwendeten Werkzeuge **unter Angabe des Verwendungszwecks in der Arbeit aufgelistet und transparent gemacht werden, z.B. in einem spezifischen Verzeichnis** (vgl. bereits Ziff. 6.2.II RL KI).

Wird mit einem Literaturverwaltungsprogramm (z. B. Zotero) gearbeitet, welches generative KI-Systeme im allgemeinen Literaturverzeichnis nachweist und kann der Verwendungszweck nicht im Verzeichnis selbst abgebildet werden, so ist sicher zu stellen, dass in der Methoden-Referenz gemäss Ziff. 3.IV der jeweilige Verwendungszweck der eingesetzten KI-Systeme detailliert dokumentiert ist (vgl. Beispiel 2 unter Ziff. 3.IV).

Zitiervorgabe³:

1. Autor des Modells (z.B. OpenAI als Autor des Large Language Models ChatGPT)
2. Datum (i.e. das Jahr in welchem ein generatives KI-System genutzt wurde (und nicht das exakte Verwendungsdatum))
3. Titel (i.e. Name des Modells, z.B. ChatGPT, mit *Kursivschrift*)
4. Version (i.e. Bezeichnung der Version oder Versionierung)
5. Quelle (i.e. URL Link)

³ Quelle: <https://apastyle.apa.org/blog/how-to-cite-chatgpt>, abgerufen am 27.10.2023.



Beispiele (Spezifisches Verzeichnis):

- Literaturverzeichnis:
[...]
 - Verzeichnis Generativer KI-Systeme:
 - OpenAI. (2023). *ChatGPT* (Version 3.5).
<https://chat.openai.com/chat>
 - Erstellung von Textvorschlägen
 - Entwicklung einer Methodologie
 - Bearbeitung von Daten
 - Erstellung von Fragebogen
 - Strukturierungshilfe für Unterteilung der Arbeit
 - Open AI. (2023). DALL-E (Version 3).
<https://openai.com/dall-e-3>
 - Bildgenerierung
 - DeepL (2023). *DeepL Translator* (Version DeepL Pro).
<https://www.deepl.com/translator>
 - Übersetzen von Textpassagen
 - Litmaps (2023). *Litmaps* (Free Version).
<https://www.litmaps.com>
 - Literaturrecherche
- III. Die **Prompts**, welche im Dialog mit den generativen KI-Systemen verwendet wurden, sowie der korrespondierende Output sind *nicht* zu deklarieren, d.h. sie sind weder als «In-Text Referenz» (Ziff. 3.I) noch als «Verzeichnis-Referenz» (Ziff. 3.2) anzugeben oder aufzulisten (vgl. als Ausnahme die Bildbezeichnung unter Ziff. 3.I).
- IV. Zur besseren Erkennbarkeit des Anteils bzw. des Ausmasses des Mitwirkens von generativen KI-Systemen an der schöpferischen Leistung eine Arbeit sowie zur Dokumentation der Qualität des eigenen wissenschaftlichen Wirkens, muss generell **in der Einleitung zur Arbeit oder im Methodenabschnitt/-teil beschrieben werden, wie generative KI-Systeme in der Arbeit verwendet wurden**. Zur Beschreibung gehören zwingend Ausführungen darüber, inwiefern generative KI-Systeme für einen Gedankenanstoss oder als Inspirationsquelle verwendet wurden. Werden zudem generative KI-Systeme wie unter Ziff. I oder II verwendet, muss dies ergänzend dokumentiert werden. Wird der Verwendungszweck im Verzeichnis über eingesetzte generative KI-Systeme spezifisch angegeben, reicht eine summarische Beschreibung in der Methoden-Referenz (Ziff. 3.II).

Beispiele (Methoden-Referenz):

- Beispiel 1 (summarisch bei spezifischem Verzeichnis):

Einleitung

[...]

Schliesslich sei erwähnt, dass generative KI-Systeme bzw. KI-Tools einerseits als Inspirationsquelle und für ein initiales Brainstorming verwendet wurden. Besonders die kritische dialogische Auseinandersetzung mit [...] und seinen Inhalten hat meine Arbeit und die Qualität meines wissenschaftlichen Wirkens bereichert. Durch das Modell konnte ich nicht nur Ideen generieren, sondern auch die Grenzen und möglichen Verzerrungen in den erzeugten Inhalten erkennen, z.B. bei dem Aspekt [...]. Dies förderte ein tieferes Verständnis für die Anwendung von Sprachtechnologien und sensibilisierte für deren kritische Reflexion in meiner Arbeit. Andererseits wurden KI-Tools u.a. für das Recherchieren von Literatur, das Strukturieren der Arbeit in Unterkapitel, das Zusammenfassen von Literatur, das Verstehen von Inhalten, das Formulieren von Texten, oder die Bildgenerierung verwendet. Es kann hierfür auf das spezifische Verzeichnis verwiesen werden.

- Beispiel 2 (detaillierend bei fehlendem spezifischem Verzeichnis bzw. bei fehlender Möglichkeit Verwendungszweck von generativen KI Systemen in Anwendung von Literaturverwaltungsprogrammen im allgemeinen Literaturverzeichnis selbst abzubilden):

Einleitung

[...]

Schliesslich sei erwähnt, dass generative KI-Systeme bzw. KI-Tools einerseits als Inspirationsquelle und für ein initiales Brainstorming verwendet wurden. Besonders die kritische dialogische Auseinandersetzung mit [...] und seinen Inhalten hat meine Arbeit und die Qualität meines wissenschaftlichen Wirkens bereichert. Durch das Modell konnte ich nicht nur Ideen generieren, sondern auch die Grenzen und möglichen Verzerrungen in den erzeugten Inhalten erkennen, z.B. bei dem Aspekt [...]. Dies förderte ein tieferes Verständnis für die Anwendung von Sprachtechnologien und sensibilisierte für deren kritische Reflexion in meiner Arbeit. Andererseits wurde für die Schärfung des Arbeitstitels [Forschungsfrage], das Strukturieren der Arbeit in Unterkapitel, das Zusammenfassen von Literatur, das Verstehen von Inhalten, sowie das Formulieren von Texten ChatGPT verwendet. Ferner wurde Litmaps für das Recherchieren von Literatur eingesetzt. Schliesslich wurde zur Bildgenerierung das generative KI-Systeme DALL-E gebraucht.

- V. **Ausgenommen von der Deklarationspflicht** ist der Einsatz generativer KI-Systeme:
- als Formulierungs- und Rechtschreibhilfe (Sprachunterstützung); oder
 - für Sprachergänzung.



4. Verantwortung

Die Verantwortung bleibt bei den Studierenden bzw. Weiterbildungsteilnehmenden. Auch bei korrekter Deklaration bleiben die Studierenden bzw. Weiterbildungsteilnehmenden verpflichtet, die Relevanz, den Wahrheitsgehalt oder die Genauigkeit sowohl der eingesetzten generativen KI-Systeme als auch des Outputs (insb. bei Literaturrecherchen und Quellen) zu überprüfen (Stichwort: Halluzinieren).

Arbeiten müssen zudem eigenständige Leistungen der Studierenden bzw. Weiterbildungsteilnehmenden sein (vgl. auch Ziff. 7 RL KI). Folglich dürfen generative KI-Systeme bei Arbeiten ausschliesslich unterstützend eingesetzt werden. Studierende bzw. Weiterbildungsteilnehmende müssen mit ihnen reflektierend, kritisch und im Ergebnis steuernd umgehen.

Schliesslich gilt es zu beachten, dass besonders schützenswerte Personendaten (z.B. Daten über religiöse Ansichten, Gesundheit, oder genetische Daten) sowie Betriebs- und Geschäftsgeheimnisse (juristischer Personen) nicht mit generativen KI-Systeme bearbeitet werden dürfen.

5. Erlassinformationen

5.1 Metadaten Erlass

Betreff	Inhalt
Erlassverantwortliche:r	Leiter:in Stab Ressort Bildung
Beschlussinstanz	Leiter:in Ressort Bildung
Themenzuordnung	2.05.00 Lehre Studium / 5.04.00 Durchführung WB
Publikationsart	Public

5.2 Erlassverlauf

Version	Beschluss	Beschlussinstanz	Inkrafttreten	Beschreibung Änderung
1.0.0	28.11.2023	Leiter:in Ressort Bildung	01.12.2023	Originalversion