

KI- ETHIKLEITLINIE LEHRE

Zum Umgang mit Künstlicher Intelligenz in der Lehre der HFT Stuttgart

01. EINLEITUNG

Der Einsatz von Künstlicher Intelligenz (KI) wird in Zukunft auf allen Ebenen an Bedeutung gewinnen. Die HFT Stuttgart sieht es als Teil ihres Bildungsauftrags an, ihre Studierenden angemessen auf diese Realität vorzubereiten. Aspekte eines mündigen Umgangs mit KI sind die kritische Reflexion ihrer Nutzung, die Auseinandersetzung mit den Zwecken, die mit ihr verfolgt werden, und das Bewusstsein für die kurz- und langfristigen Auswirkungen, die ihr Einsatz mit sich bringt.

Die *multi-purpose* Technologie Künstliche Intelligenz birgt große Potentiale für viele Bereiche des täglichen Lebens. Dies gilt auch für den Lehrbetrieb der Hochschulen. Die HFT Stuttgart erkundet stetig neue technische Möglichkeiten, ihre Lehre weiter zu verbessern und strebt deshalb die erfolgreiche Integration von KI in den Hochschulbetrieb an. Dies bedeutet unter anderem, dass die Nutzung von KI an der HFT Stuttgart im Einklang mit ihren Kernwerten stattfindet. Die vorliegende Leitlinie dient als Orientierungshilfe für den wertebasierten Einsatz von KI in der Lehre der HFT Stuttgart. Sie richtet sich damit sowohl an Lehrende und Mitarbeitende der HFT Stuttgart, die direkt oder indirekt in die Lehre involviert sind, sowie an Studierende der HFT Stuttgart.

Die zugrundeliegenden Werte und Kernüberzeugungen dieser Leitlinie sind im Ethikkodex der HFT Stuttgart niedergelegt. Aufgrund der Dynamik, die der technologischen Entwicklung innewohnt, und der Unvorhersehbarkeit neuer KI-Tools muss diese Leitlinie im Sinne eines lebenden Dokuments kontinuierlich weiterentwickelt werden, um angemessen auf künftige Bedarfe einzugehen und sicherzustellen, dass den im Ethikkodex niedergelegten Werten im Kontext aktueller Entwicklungen Rechnung getragen wird. Um Nachvollziehbarkeit zu gewährleisten und die laufende Aktualisierung des Dokuments zu vereinfachen, sind jedem Sinnabschnitt die jeweiligen Leitmotive vorangestellt.

02. GRUNDLEGENDES

Definition: KÜNSTLICHE INTELLIGENZ

Künstliche Intelligenz ist ein Teilgebiet der Informatik, das sich mit der Automatisierung von intelligentem Verhalten befasst. Dabei ist weder festgelegt, was »intelligent« bedeutet, noch welche Technik zum Einsatz kommt. Eine der Grundlagen der modernen Künstlichen Intelligenz ist das Maschinelle Lernen. Weitere wichtige Methoden sind logisches Schließen auf symbolischem Wissen, Wissensrepräsentation oder Planungsverfahren. [...]

<https://www.ki.nrw/ki-schlüsselbegriffe/#22>

2.1. Kritische Reflexion und Nutzung neuer Potenziale

Der mündige Umgang mit KI erfordert die kritische Reflexion ihres Einsatzes. Während eine grundsätzliche Ablehnung Potentiale ungenutzt lässt, unterschätzt eine unkritische Technologiebegeisterung die Risiken von KI.

1. Die HFT Stuttgart empfiehlt Lehrpersonen nachdrücklich, die Chancen und Risiken von KI sowie die ethischen Herausforderungen, die mit ihr im jeweiligen Anwendungskontext einhergehen, aktiv in Lehrveranstaltungen mit Bezug zu KI mitzudenken und in den Unterricht zu integrieren. Dazu gehört auch, die Potentiale und Grenzen konkreter KI-Anwendungen realistisch einzuschätzen.
2. Die Nutzung von KI-Anwendungen kann dazu beitragen, Lehrveranstaltungen ansprechender und gewinnbringender für Studierende zu gestalten. Lehrpersonen, die KI-Anwendungen in einer Lehrveranstaltung einsetzen möchten, sind angehalten, diesen Einsatz kritisch zu hinterfragen, idealerweise gemeinsam mit den Teilnehmenden der Veranstaltung.

Weiterführendes Material



Übersicht über Chancen und Risiken von Künstlicher Intelligenz:

<https://www.europarl.europa.eu/topics/de/article/20200918STO87404/kunstliche-intelligenz-chancen-und-risiken>

Materialverzeichnis zu existenziellen Risiken von Künstlicher Intelligenz:

<https://futureoflife.org/resource/introductory-resources-on-ai-risks/>

Ressourcen für das Selbststudium:

- KI-Campus:
<https://ki-campus.org/>
- Deutsche Gesellschaft für Hochschuldidaktik
<https://www.dghd.de/praxis/veranstaltungskalender/dghd-themenreihe-ki-in-der-hochschullehre/>

2.2. Lebenslanges Lernen und Aktualität

Die HFT Stuttgart ist bemüht, ihren Studierenden innovative und aktuelle Lehrangebote zu unterbreiten, um hochwertige Bildung sicherzustellen.

1. Die HFT Stuttgart empfiehlt Lehrpersonen, sich selbstständig und laufend über Anwendungsmöglichkeiten von KI in ihren jeweiligen Fachbereichen zu informieren und aktuelle Entwicklungen in ihre Lehrveranstaltungen zu integrieren.
2. Die HFT Stuttgart ist zudem bestrebt, ihren Dozierenden regelmäßig Fort- und Weiterbildungsangebote im Kontext von KI zu unterbreiten.

Weiterführendes Material



Übersicht über Anwendungsmöglichkeiten von Künstlicher Intelligenz:

<https://research.aimultiple.com/generative-ai-applications/>

Fachübergreifende Anwendungsmöglichkeiten von ChatGPT in der Lehre:

[https://wiki.hft-](https://wiki.hft-stuttgart.de/pages/viewpage.action?spaceKey=NNHS&title=Handreichung+ChatGPT)

[stuttgart.de/pages/viewpage.action?spaceKey=NNHS&title=Handreichung+ChatGPT](https://wiki.hft-stuttgart.de/pages/viewpage.action?spaceKey=NNHS&title=Handreichung+ChatGPT)
[#HandreichungChatGPT-Wiel%C3%A4sstichChatGPTinStudiumundLehreeinsetzen?](#)

2.3. EU KI-Verordnung

Die KI-Verordnung der EU stellt den rechtlichen Rahmen für die Nutzung künstlicher Intelligenz dar und gilt auch für den Hochschulkontext.

1. Zum 01. August 2024 ist die KI-Verordnung¹ (KI-VO) der EU in Kraft getreten.² Die vorliegende Leitlinie erhebt nicht den Anspruch, die dort aufgestellten Regelungen zu ersetzen. Für alle Entscheidungen, die in den Anwendungsbereich der KI-VO fallen, ist die KI-VO maßgeblich, die vorliegende Leitlinie gilt demgegenüber nachrangig.
2. Die KI-VO verfolgt einen risikobasierten Ansatz und unterteilt KI-Systeme in vier Kategorien.
 - Verbotene KI-Praktiken
 - Hochrisiko-KI-Systeme
 - KI-Systeme mit beschränktem Risiko
 - KI-Systeme mit geringem RisikoJe nachdem, in welche Kategorie ein KI-System in einem bestimmten Anwendungskontext fällt, können daraus verschiedene Pflichten und Vorgaben für Anbietende, Betreibende und Endnutzende resultieren. Lehrende und Mitarbeitende sind selbst in der Verantwortung, sicherzustellen, dass der Einsatz von KI-Systemen in Lehrveranstaltungen der HFT Stuttgart nicht der VO widerspricht.³
3. Auch im Bildungskontext klassifiziert die KI-VO bestimmte KI-Systeme und KI-Praktiken als Hochrisiko-KI-Systeme.⁴ Hochrisiko-KI-Systeme im Bildungskontext sind:
 - KI-Systeme, die dazu dienen, den Zugang oder die Zulassung zu Einrichtungen der allgemeinen und beruflichen Bildung auf allen Ebenen zu bestimmen;
 - KI-Systeme, die zur Bewertung von Lernergebnissen eingesetzt werden sollen;
 - KI-Systeme, die zur Bewertung des angemessenen Bildungsniveaus, das eine Person erhalten wird oder zu dem sie Zugang haben wird, eingesetzt werden sollen;
 - KI-Systeme, die zur Überwachung und Erkennung von unzulässigem Verhalten von Schülern oder Studierenden bei Prüfungen eingesetzt werden sollen.

Die Regelungen, die beim Einsatz dieser Systeme gelten (sowie Ausnahmen davon) finden sich in der KI-VO. Weitere Informationen sind in der geplanten „Richtlinie zur Nutzung von Künstlicher Intelligenz (KI)“ der HFT Stuttgart enthalten (derzeit noch in der Entwicklung, Stand 01.12.2025).

4. Die KI-VO selbst verweist auf die Rolle, die ethischen Überlegungen – insbesondere bei der Erstellung von Verhaltenskodizes – zukommt.⁵ In diesem Kontext werden die sieben Prinzipien für ethische KI der Expertenkommission der EU aufgeführt.
 - Menschliches Handeln und Aufsicht
 - Technische Robustheit und Sicherheit
 - Privatsphäre und Daten-Governance
 - Transparenz

¹ <https://artificialintelligenceact.eu/>.

² <https://artificialintelligenceact.eu/de/implementation-timeline/>.

³ Vgl. hierzu <https://artificialintelligenceact.eu/article/26/> und <https://artificialintelligenceact.eu/article/13/>.

⁴ <https://artificialintelligenceact.eu/annex/3/>, <https://artificialintelligenceact.eu/recital/56/>.

⁵ <https://artificialintelligenceact.eu/recital/27/>, <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.

- Vielfalt, Nicht-Diskriminierung und Fairness
- Soziales und ökologisches Wohlergehen
- Rechenschaftspflicht.

Die vorliegende Leitlinie ist vor dem Hintergrund dieser Gestaltungsprinzipien entstanden und erhebt den Anspruch, ihnen Rechnung zu tragen. Ungeachtet dessen setzt die Umsetzung der Leitlinie sowie der oben genannten Prinzipien das eigenverantwortliche und aktive Mitwirken aller Adressat:innen dieser Leitlinie voraus.

Weiterführendes Material

Pflichten aus der KI-VO für **Anbieter** von KI-Systemen:



Dokumentation & Kontrolle

- Risikomanagement, Art. 9 KI-VO
- Daten-Governance, Art. 10 KI-VO
- Technische Dokumentation, Art. 11, 18 KI-VO
- Aufzeichnung und Aufbewahrung, Art. 12, 19 KI-VO
- Transparenzpflichten, Art. 13 KI-VO
- Menschliche Aufsicht, Art. 14 KI-VO
- Genauigkeit, Robustheit, Cybersicherheit, Art. 15 KI-VO
- Qualitätsmanagement, Art. 17 KI-VO
- Korrekturmaßnahmen, Art. 20 KI-VO
- Zusammenarbeit mit Behörden, Art. 21 KI-VO
- Beobachtungspflicht, Art. 72 KI-VO

Schulungspflichten

- Training der KI-Kompetenz, Art. 4 KI-VO

Sonstige Pflichten

- Konformitätsbewertung, Art. 43 KI-VO
- Transparenzpflichten, Art. 50 KI-VO („Kennzeichnung“, Sonstige KI)
- Sonstige Rechtspflichten abseits der KI-VO, insbesondere DSGVO

Pflichten aus der KI-VO für **Betreiber** von KI-Systemen:

Dokumentation & Kontrolle

- Technische und organisatorische Maßnahmen, Art. 26 I KI-VO
- Menschliche Aufsicht, Art. 26 II KI-VO
- Kontrolle der Eingabedaten, Art. 26 IV KI-VO
- Fortlaufende Überwachung, Art. 26 V, VII KI-VO, ggf. Information des Anbieters, der *Arbeitnehmenden und ggf. Aussetzung des Betriebs*
- Grundrechte-Folgenabschätzung, Art. 27 KI-VO

Anbieterpflichten

- Wesentliche Änderung oder Änderung der Zweckbestimmung, Art. 25 KI-VO

Schulungspflichten

- Training der KI-Kompetenz, Art. 4 KI-VO
- Training von Aufsichtspersonen, Art. 26 II KI-VO

Sonstige Pflichten

- Transparenzpflichten, Art. 26 XI, 50 KI-VO („Kennzeichnung“)
- Sonstige Rechtspflichten abseits der KI-VO, insbesondere DSGVO

Handreichungen zu Rechtsfragen rund um KI im Kontext der Hochschullehre:

<https://www.hnd-bw.de/projekte/bwdigirecht/archiv-handreichungen-bwdr-2-2/>

2.4. Datenschutz

Datenschutz ist generell – aber auch insbesondere in Hinblick auf KI-Tools – essentiell.

1. Nutzende von KI-Tools sind genau wie in anderen Kontexten an die datenschutzrechtlichen Vorgaben gebunden. Dementsprechend gilt es auch bei der Nutzung von KI-Tools die Datenschutzgrundsätze einzuhalten. Zu diesen Grundsätzen gehören u.a. die sparsamen Verarbeitungen personenbezogener Daten, die Erfüllung von Transparenzpflichten, das

Vorliegen einer rechtlichen Grundlage sowie der Einsatz ausreichender Sicherheitsmaßnahmen. Zudem sind die Schutzziele des Datenschutzes (Verfügbarkeit, Vertraulichkeit und Integrität) bei der Nutzung von KI zu gewährleisten.

2. Die Datenschutzvorgaben werden in der geplanten „Richtlinie zur Nutzung von Künstlicher Intelligenz (KI)“ der HFT Stuttgart beschrieben (derzeit noch in der Entwicklung, Stand 01.12.2025). Bei Unsicherheiten und Fragen zum Datenschutz ist die:der Datenschutzbeauftragte der HFT Stuttgart zu konsultieren, vgl. weiterführendes Material.

Weiterführendes Material



Datenschutz an der HFT Stuttgart:

<https://wiki.hft-stuttgart.de/pages/viewpage.action?spaceKey=KoordinationDS75&title=Datenschutz>

Kontakt Daten Datenschutzbeauftragter:

Maximilian Musch

+49 (0)7542 949 2102

datenschutz@hft-stuttgart.de

2.5. Umweltschutz und Nachhaltigkeit

KI-Systeme sollten unter Berücksichtigung des damit verbundenen Energieverbrauchs ausgewählt werden.

1. KI-Systeme basieren oft auf großen Datenmengen und hoher Rechenleistung. Damit geht ein entsprechend hoher Energieverbrauch einher, der oftmals einen hohen Treibhausgasausstoß beinhaltet. Die HFT Stuttgart sieht sich unter anderem dem Umweltschutz verpflichtet, der unter den zuvor genannten Aspekten einen sparsamen Umgang mit KI-Systemen nahelegt.
2. Gleichzeitig beansprucht die HFT Stuttgart für sich, ihren Studierenden hochwertige Bildungsangebote zu unterbreiten, die aktuellen Entwicklungen, insbesondere in technischen Bereichen, Rechnung tragen. Eine allgemeingültige Lösung zugunsten oder entgegen des Einsatzes von KI kann aufgrund der Komplexitäten des Themas nicht gegeben werden. Vielmehr muss der Einsatz von KI an der HFT Stuttgart kritisch und in Hinblick auf dieses Spannungsfeld abgewogen werden.
3. Lehrpersonen, die erwägen, KI-Systeme in ihren Lehrveranstaltungen einzusetzen, wird empfohlen, den Energieverbrauch (sofern bekannt bzw. recherchierbar) verschiedener infrage kommender KI-Systeme in ihre Abwägung miteinzubeziehen.
4. Lehrpersonen, die KI-Systeme in ihren Lehrveranstaltungen einsetzen wollen, wird empfohlen, über den damit einhergehenden Energieverbrauch kritisch mit ihren Studierenden zu reflektieren.

03. EINSATZ VON KÜNSTLICHER INTELLIGENZ IN LEHRVERANSTALTUNGEN DER HFT STUTTGART

3.1. Sicherheit

Der Einsatz von KI in Lehrveranstaltungen darf nicht zu Lasten der Studierenden erfolgen.

1. Lehrpersonen gestalten den Einsatz von KI in ihren Lehrveranstaltungen so, dass er gegenüber jeder:m einzelnen Studierenden vertreten werden kann.
2. Dies bedeutet unter anderem, dass der Einsatz von KI-Tools in Lehrveranstaltungen nur zulässig ist, wenn Gefährdungen der physischen und psychischen Gesundheit (vgl. **BEISPIEL 1**) der Teilnehmenden so gut wie vernünftig möglich vermieden werden, ihre Persönlichkeitsrechte gewahrt bleiben und personenbezogene Daten ausschließlich im Rahmen der Grundsätze der DSGVO verarbeitet werden.
3. Besondere Sorgfalt ist in Bezug auf KI-Tools zu wahren, die Studierende auch außerhalb von Lehrveranstaltungen nutzen können (oder: nutzen sollen). Lehrpersonen sind hier gehalten, nach bestem Wissen und Gewissen erwartbare Fehlgebrauch, der zulasten der physischen oder psychischen Gesundheit ihrer Studierenden geht, proaktiv entgegenzuwirken.

4. KI-Tools, bei denen nicht hinreichend sichergestellt werden kann, dass ihre Nutzung außerhalb von Lehrveranstaltungen ohne eine Gefährdung der physischen und psychischen Gesundheit von Studierenden erfolgt, dürfen Studierenden nicht zur Nutzung außerhalb von Lehrveranstaltungen überlassen werden, es sei denn, eine solche Gefährdung der physischen oder psychischen Gesundheit setzt mutwilligen Missbrauch der Anwendung voraus.



BEISPIEL 1:

Lehrperson A möchte in einer Lehrveranstaltung ein KI-Tool vorführen, dass auf der Basis von Prompts selbstständig kurze Videos generiert. Um das psychische Wohlbefinden der Teilnehmenden der Lehrveranstaltung nicht zu gefährden, lässt Lehrperson A keine Videos generieren, in denen gewaltsame Handlungen dargestellt werden.

3.2. Zugangsgerechtigkeit und Freiwilligkeit

Studierende haben einen Anspruch auf Gleichbehandlung. Dieser erstreckt sich prinzipiell auch auf den Zugang zu Lernmaterialien.

1. Sollen die Teilnehmenden einer Veranstaltung im Rahmen der Veranstaltung ein KI-Tool nutzen, muss der Zugang dazu durch die Lehrperson sichergestellt werden (bspw. über HFT-Lizenzen), sofern diese nicht frei verfügbar sind.
2. Die Nutzung von KI-Tools darf nur zur Voraussetzung der Teilnahme an einer Veranstaltung gemacht werden, wenn die Lernziele der Veranstaltung nicht auf anderem Wege oder nur mit deutlich höherem Zeitaufwand erreicht werden können *und* wenn diese Voraussetzung im Vorfeld kommuniziert wurde, bspw. im Rahmen der Modulbeschreibung, der Kursbeschreibung im LSF oder des ersten Veranstaltungstermins der entsprechenden Lehrveranstaltung.
Erfordert die Nutzung eines KI-Tools die Preisgabe persönlicher Daten, so gilt ferner, dass sie nur zur Voraussetzung der Teilnahme an einer Veranstaltung gemacht werden darf, wenn persönliche Daten ausschließlich an Personen oder Organisationen preisgegeben werden, die den Bestimmungen der DSGVO Rechnung tragen *und* wenn die Notwendigkeit der Preisgabe persönlicher Daten ebenfalls im Vorfeld kommuniziert wurde.
3. Studierende dürfen nicht genötigt werden, gegen ihren Willen persönliche Daten von sich preiszugeben (etwa indem sie einen Account unter ihrem Klarnamen erstellen), um Zugang zu einem KI-Tool zu erhalten.
4. Sofern die Teilnehmenden einer Lehrveranstaltung im Rahmen der Veranstaltung bestimmte KI-Tools nutzen sollen, muss die Lehrperson sicherstellen, dass auch Teilnehmende der Veranstaltung mit körperlichen oder geistigen Einschränkungen die Möglichkeit haben, diese KI-Tools effektiv zu nutzen. Lehrpersonen sollten sich deshalb rechtzeitig über alternative Nutzungsmöglichkeiten, KI-Tools oder Lernmöglichkeiten informieren.

3.3. Transparenz

Aus Gründen der Nachvollziehbarkeit steht der Einsatz von KI generell unter dem Gebot der Transparenz.

1. Der Einsatz von KI in Lehrveranstaltungen muss transparent erfolgen und den Teilnehmenden rechtzeitig kommuniziert werden, bspw. im Rahmen der Modulbeschreibung, der Kursbeschreibung im LSF oder des ersten Veranstaltungstermins der entsprechenden Lehrveranstaltung. Dies beinhaltet die explizite Benennung von KI-Systemen als solche, es sei denn, ein solches System ist für eine umsichtige, vernünftige Person als solches ohne weiteres erkennbar.
2. Der transparente Umgang mit KI beinhaltet auch die Kenntlichmachung von KI-generierten Inhalten und die entsprechende Kennzeichnung von sonstiger unterstützender KI-Nutzung (bspw. automatisierte Übersetzungen ganzer Textteile).

3.4. Diskriminierende KI-Systeme

Diskriminierende Biases in KI-Systemen verstetigen schädliche Vorurteile. KI-Tools, die mit historisch gewachsenen Daten trainiert wurden, Muster in diesen Daten erkennen und auf dieser Basis Vorhersagen treffen, sind inhärent anfällig für diskriminierende Biases. Lehrpersonen sind in der Pflicht, diesem Sachverhalt mit angemessenen Maßnahmen entgegenzuwirken.

Definition: DISKRIMINIERENDER BIAS

KI-Tools können diskriminierende Biases aufweisen. Als „Bias“ werden systematische Fehleinschätzungen bezeichnet. Im Kontext von Künstlicher Intelligenz spricht man von „biased AI“, wenn bspw. Gesichtserkennungssoftware bestimmte Menschengruppen besser oder schlechter erkennt als andere.

Ein Bias ist diskriminierend, wenn er ungerechtfertigt ist und Bezug auf bestimmte Charakteristika nimmt. Welche Charakteristika genau in welchen Kontexten als diskriminierend bezeichnet werden können, ist umstritten. Orientierung bietet bspw. Artikel 14 der Europäischen Menschenrechtskonvention. Anknüpfend an das obige Beispiel wäre der Bias der Gesichtserkennungssoftware diskriminierend, wenn die besser oder schlechter erkannte Personengruppe etwa aufgrund ihrer Hautfarbe oder ihres Geschlechts besser oder schlechter erkannt würde. Umgekehrt wäre ein Bias vermutlich nicht diskriminierend, wenn die Gesichtserkennungssoftware Menschen besser erkennt, die eine Hand in die Kamera halten, als solche, die dies nicht tun.

1. Lehrpersonen, die ein KI-Tool in einer Lehrveranstaltung nutzen wollen, sind angehalten, auf das Auftreten diskriminierender Biases in KI-Tools zu achten und gegebenenfalls in der Lehrveranstaltung zu thematisieren. Hierfür kann der Foliensatz „AI Bias“ verwendet werden, vgl. weiterführendes Material.
2. Ist bekannt, dass ein gegebenes KI-Tool bestimmte diskriminierende Biases aufweist, so sollte dieses KI-Tool nur verwendet werden, wenn gewichtige Gründe für seinen Einsatz sprechen, etwa ein Mangel an alternativen Tools in Kombination mit einem sehr hohen praktischen Nutzen des Tools oder wenn davon ausgegangen werden kann, dass der diskriminierende Bias im Rahmen der Lehrveranstaltung nicht unmittelbar relevant wird, vgl. **BEISPIEL 2**.

**BEISPIEL 2:**

Lehrperson B will in einer Lehrveranstaltung ein bildgenerierendes KI-Tool einsetzen. B ist bekannt, dass das KI-Tool bei der Generierung von Bildern von Menschen kulturelle Stereotype bedient. Da die Studierenden das Tool allerdings ausschließlich für die Erstellung geometrischer Figuren nutzen sollen, kann B davon ausgehen, dass der beschriebene Bias in der Lehrveranstaltung nicht relevant wird.

3. Im Fall eines solchen Einsatzes sind Lehrpersonen angehalten, diskriminierende Biases des Tools proaktiv in der Lehrveranstaltung zu thematisieren und das Auftreten diskriminierender Biases des KI-Tools in der Lehrveranstaltung zu thematisieren und kritisch zu reflektieren. Bei Unsicherheiten und Rückfragen wenden Sie sich an den:die Gleichstellungsbeauftragte:n der HFT Stuttgart, vgl. weiterführendes Material.

Weiterführendes Material

Foliensatz „AI Bias“:

>>>Wird mit Verabschiedung der Leitlinie im HFT-Wiki veröffentlicht<<<



Chancengleichheit an der HFT Stuttgart:

<https://www.hft-stuttgart.de/hft/campuskultur/chancengleichheit/gleichstellung>

Kontaktdaten Gleichstellungsbeauftragte:

Christine Kappei

+49 711 8926 2590

christine.kappei@hft-stuttgart.de

3.5. Nutzung von KI-Systemen im Rahmen von Benotungen

Studierende haben einen Anspruch darauf, dass ihre Leistungen von einem Menschen bewertet werden. Lehrpersonen sind und bleiben für die Bewertung von Prüfungsleistungen verantwortlich.

1. Die abschließende Benotung studentischer Prüfungsleistungen durch KI-Systeme ist untersagt.
2. Sofern KI-Systeme genutzt werden, um Anhaltspunkte für die abschließende Bewertung einer für die Modulnote relevanten Leistung zu generieren, müssen Art und Umfang dieses Einsatzes den Studierenden offengelegt werden (bspw. im Rahmen der Kursbeschreibung oder im zugehörigen Moodle-Kurs). Diese KI-generierten Anhaltspunkte müssen stets von einem Menschen auf ihre Richtigkeit geprüft werden. Dies beinhaltet alle KI-Systeme, die studentische Arbeitsergebnisse oder Teile von studentischen Arbeitsergebnissen analysieren und eine qualitative oder quantitative Rückmeldung dazu generieren.
3. Für Studierende muss die Möglichkeit bestehen, die Bewertungen des KI-Systems anzufechten und korrigieren zu lassen. Studierende sind auf die Anfechtbarkeit der Bewertungen des KI-Tools aufmerksam zu machen.
4. Manche KI-Tools sind besonders anfällig dafür, existierende Vorurteile und Normalitätsvorstellungen aus ihren Trainingsdaten zu übernehmen, vgl. **3.4 Diskriminierende Biases**. Lehrpersonen, die KI-Tools im Rahmen der Leistungsbewertung nutzen, müssen nach bestem Wissen und Gewissen sicherstellen, dass daraus keine unfairen Nachteile für einzelne Studierende resultieren.

3.6. Nutzung von KI-Systemen im Rahmen von sonstigen Beurteilungen

Studierende haben einen Anspruch darauf, dass Bewertungen ihrer Arbeitsergebnisse transparent, nachvollziehbar und kriterienbasiert sind, auch wenn sie keinen unmittelbaren Einfluss auf die Benotung haben.

1. Der Einsatz von KI-Systemen zur Erstellung von Bewertungen studentischer Arbeitsergebnisse, die nicht relevant für die abschließende Leistungsbewertung sind, müssen Studierenden offengelegt werden. Dies beinhaltet alle KI-Systeme, die studentische Arbeitsergebnisse oder Teile von studentischen Arbeitsergebnissen analysieren und eine qualitative oder quantitative Rückmeldung dazu generieren.
2. Für Studierende muss die Möglichkeit bestehen, die Bewertungen des KI-Systems anzufechten. Studierende sind auf die Anfechtbarkeit der Bewertungen des KI-Tools aufmerksam zu machen, vgl. **BEISPIEL 3**.



BEISPIEL 3:

Lehrperson C will in einer Lehrveranstaltung ein Sprachmodell einsetzen, das den Studierenden Feedback zu Programmieraufgaben gibt. Obwohl dieses Feedback nicht in die abschließende Leistungsbewertung eingeht, muss Lehrperson C ihren Studierenden die Möglichkeit geben, die Bewertungen des KI-Tools anzufechten. Studierende dieser Lehrveranstaltung erhalten deshalb die Möglichkeit, die Bewertungen des KI-Tools durch Lehrperson C prüfen und erläutern zu lassen.

04. GENERATIVE KI UND PRÜFUNGSLEISTUNGEN

4.1. Generative KI

Der Einsatz von KI-Tools (insbesondere: generativer KI-Tools) im Rahmen von Prüfungsleistungen ist seit dem Jahr 2022 eine viel diskutierte Herausforderung für Bildungseinrichtungen. Generative KI ist in der Lage, auf Basis der Daten, mit denen sie trainiert wurde, Neues zu generieren – je nach KI kann es sich dabei bspw. um Texte, Bilder oder Audio-Dateien handeln. Am Beispiel textgenerativer KI lässt sich die Funktionsweise generativer KI veranschaulichen: Durch das Erkennen von Mustern in den Texten, mit denen die KI trainiert wurde, wird ein Sprachmodell erstellt. Dieses Sprachmodell ermittelt für jede Eingabe eine Ausgabe, indem es diejenigen Worte aneinanderreihet, die mit der höchsten Wahrscheinlichkeit die „passende“ Ausgabe darstellen. Diese ungefähre Funktionsweise findet sich ebenfalls in den Textvorschlägen, die ein Mobiltelefon anzeigt, wenn man beginnt, eine Nachricht zu verfassen. Unterschiede bestehen in der Komplexität, der Menge der Trainingsdaten und darin, was genau eine „passende“ Ausgabe konstituiert – bei den Textvorschlägen des Mobiltelefons ist es dasjenige Wort, das den angefangenen Satz am sinnvollsten fortführt, bei generativer KI wie ChatGPT hingegen, welche Wortfolge eine hilfreiche und akkurate Reaktion auf die Eingabe darstellt.

Moderne Sprachmodelle sind in der Lage, sprachlich gelungene und wohlstrukturierte Texte zu erzeugen. Zusätzlich ist die Wissensbreite der KIs aufgrund der schieren Menge an unterschiedlichen Trainingsdaten sehr hoch. Für das wissenschaftliche Arbeiten sind diese Fähigkeiten wertvoll. Die *Handreichung ChatGPT* der HFT Stuttgart (vgl. [Weiterführendes Material](#)) enthält einen Überblick über Anwendungsmöglichkeiten von KI-Tools wie ChatGPT für das wissenschaftliche Arbeiten im Studium.

4.2. Herausforderungen bei der Nutzung generativer KI

Aufgrund der zwischenzeitlich hohen sprachlichen Qualität der Texte, die moderne Sprachmodelle generieren und der thematischen Vielfalt, die in den Trainingsdaten abgebildet ist, entsteht leicht der Eindruck, dass es sich bei KIs wie ChatGPT um universalgebildete Expert:innen handelt. In der Realität unterlaufen generativen KIs häufig Flüchtigkeitsfehler, die Detailtiefe ist rasch ausgeschöpft oder das Sprachmodell „halluziniert“ plausibel scheinende „Fakten“. Im Kontext des wissenschaftlichen Arbeitens ist dies beispielsweise bei erfundenen Quellenangaben fatal. Aber auch grundsätzlichere Probleme sind mit dem Einsatz generativer KI-Tools im Kontext des wissenschaftlichen Arbeitens verbunden:

- **Verlust wissenschaftlicher Schlüsselkompetenzen**
Eine der größten Sorgen, die generative KI-Systeme insbesondere im Kontext der Hochschule begleitet, ist der Verlust (oder Nichterwerb) wissenschaftlicher Schlüsselkompetenzen durch die Nutzung von KI-Tools („Deskilling“). Wer ständig KI-Tools nutzt, um systematisch bestimmte Arbeitsschritte auszulagern, lernt nicht, diese Arbeitsschritte ohne KI zu bewältigen. Diesem Prinzip folgen auch andere technologische Neuerungen – beispielsweise war der Prozess der Literaturrecherche vor der Erfindung von Suchmaschinen ein anderer, als er es heute ist. Generative KI-Tools, insbesondere solche, die Text generieren, stellen jedoch eine neue Dimension maschineller Unterstützung dar. Während Suchmaschinen zwar etablierte Arbeitsschritte stark vereinfachen und verbessern, lassen sich mithilfe von Sprachmodellen nahezu alle Arbeitsschritte für die Erstellung akademischer Arbeiten vollständig auslagern, von der Themenfindung und der Entwicklung einer Forschungsfrage, über die Recherche relevanter Beiträge, deren Lektüre und Aufbereitung, bis hin zur Strukturierung und Ausarbeitung eines wissenschaftlichen Textes.

Dabei besteht die Versuchung, gerade in Aspekten wissenschaftlichen Arbeitens, die dem:der Nutzenden noch nicht gut gelingen, auf Sprachmodelle zurückzugreifen. In den ersten Semestern eines Hochschulstudiums ist die Qualität der Ergebnisse von Modellen wie ChatGPT vermutlich ausreichend für eine positive Bewertung. Wenn die Anforderungen des Studiums allerdings die Fähigkeiten der KI übersteigen, stehen diejenigen vor Problemen, die sich bestimmte Kompetenzen nicht angeeignet, sondern stattdessen auf eine KI zurückgegriffen haben. Dabei lassen sich generative KI-Systeme auf allen Ebenen wissenschaftlichen Arbeitens durchaus einsetzen – allerdings nur unter der Voraussetzung, dass der:die Nutzende über die erforderlichen Grundfertigkeiten verfügt, um die Ausgaben der KI kritisch bewerten zu können. Ein gewisses Maß an eigener Kompetenz in der jeweiligen Teilaufgabe ist dafür unerlässlich. Autor:innen bleiben auch bei Einsatz von KI für

die Inhalte wissenschaftlicher Ausarbeitungen verantwortlich und müssen die Einhaltung der guten wissenschaftlichen Praxis sicherstellen.

- **Rechtliches**
Im Bereich des Urheberrechts sind viele Fragen in Hinblick auf die zulässige Nutzung der Ausgaben generativer KI offen. Neben urheberrechtlichen Aspekten gibt es auch datenschutzrechtliche Bedenken bei der Nutzung von generativer KI. Die Eingabe personenbezogener Daten sollte unbedingt vermieden werden. Spiegelbildlich sollten alle Ausgaben daraufhin geprüft werden, ob sie personenbezogene Daten enthalten. Zu rechtlichen Aspekten Künstlicher Intelligenz im Kontext der Hochschullehre vgl. Weiterführendes Material bei 2.3.
- **Technische Abhängigkeit**
Die bekanntesten und besten KI-Tools sind Eigentum von privaten Akteur:innen. Das kontinuierliche Sichverlassen auf KI-Tools und damit verbundenes *Deskilling* gehen mit einer erheblichen Machtkonzentration bei diesen Akteur:innen einher. Ihre Privatheit entzieht sie direkter demokratischer Kontrolle.
- **Verengung des Erkenntniskorridors**
Wie oben beschrieben, erstellen Sprachmodelle Prognosen auf der Basis statistischer Zusammenhänge. So versuchen Chatbots, ihre Antworten auf bestimmte Parameter hin (bspw. möglichst hilfreich zu sein) statistisch zu optimieren. Wenn Menschen generative KI nutzen, um die damit generierten Outputs zu veröffentlichen (bspw. im Zuge wissenschaftlicher Arbeit), können diese Veröffentlichungen wiederum neue Daten für das KI-Tool darstellen, die das statistische Normal mitdefinieren. Je länger sich dieser Prozess fortsetzt, umso enger wird das Spektrum, in dem sich die Ausgaben der KI befinden. Wenn KI eine maßgebliche Rolle für bestimmte Teilbereiche des wissenschaftlichen Arbeitens spielt, besteht deshalb die Gefahr einer Verengung des Erkenntniskorridors.

4.3. Prüfungsleistungen

Trotz der genannten Herausforderungen stellt generative KI eine wichtige mögliche Unterstützung beim wissenschaftlichen Arbeiten dar. Zur guten Praxis wissenschaftlichen Arbeitens gehört es allerdings auch, dass Wissenschaftler:innen offenlegen, welche Quellen und Hilfsmittel in die Erstellung eines Forschungsbeitrags eingegangen sind. Derzeit ist jedoch noch umstritten, welche Nutzung welcher KI-Tools in welcher Detailtiefe dokumentiert und als Anhang mit eingereicht werden muss. Hierbei liegt es nahe, dass die diesbezüglichen Anforderungen zwischen verschiedenen Fachbereichen variieren. Solange keine hochschulweit einheitlichen Zitierregeln für den Einsatz generativer KI in Prüfungen existieren, sind die Fakultäten der HFT Stuttgart aufgefordert, entsprechende Regeln aufzustellen und ihren Studierenden zu kommunizieren. Als Anregung empfehlen sich die diesbezüglichen Festlegungen des Fachbereichs Wirtschaft der HFT Stuttgart (Merkblatt zur Anfertigung wissenschaftlicher Arbeiten, S. 18-21), die „Rules for Tools“ der Pädagogischen Hochschule Heidelberg und den Leitfaden „Aus KI zitieren“ der Universität Basel (https://dslw.philhist.unibas.ch/fileadmin/user_upload/dslw/Dokumente/MA-Studium/MSG_Sprache_und_Kommunikation/Leitfaden_KI_De_Eng_.pdf).

Grundsätzlich unterliegt die Verwendung von generativer KI bei Prüfungsleistungen folgenden Kernprinzipien:

- **Eigenverantwortung**
Studierende sind selbst für die Inhalte der Arbeiten verantwortlich, die sie als Prüfungsleistungen einreichen. Eine selbstständige Nachprüfung der Aussagen großer Sprachmodelle ist insofern unerlässlich. Fehlerhaft übernommene KI-generierte Inhalte, wie Falschaussagen und nicht-existente Literaturquellen, können zum Nichtbestehen der Prüfung führen. Grundsätzlich empfiehlt es sich, die geplante Nutzung von KI-Tools in Prüfungsleistungen mit der Lehrperson abzuklären.
- **Transparenz**
Die Nutzung von KI-Tools muss gemäß der Regelungen des jeweiligen Fakultät dokumentiert und offengelegt werden. Für Veranstaltungen, die sich nicht einer einzigen Fakultät zuordnen lassen (bspw. in Studiengängen, die von mehreren Fakultäten oder Hochschulen betrieben werden), sind die verantwortlichen Lehrpersonen bzw. Koordinator:innen in der Pflicht, Regelungen für den Umgang mit KI in Prüfungsleistungen festzulegen und den betroffenen Studierenden zu kommunizieren. Die nicht gekennzeichnete Übernahme von Ergebnissen generativer KI stellt einen Täuschungsversuch dar, der zum Nichtbestehen der Prüfung führen kann (vgl. § 14 Abs. 5 ASPO). Derzeit gibt es noch keine zuverlässige Möglichkeit, einen durch KI generierten Text als solchen zu identifizieren. Die HFT Stuttgart

behält sich jedoch vor, eingereichte Prüfungsleistungen zu einem späteren Zeitpunkt daraufhin zu prüfen, ob KI-Tools ungekennzeichnet verwendet wurden. In einem solchen Fall können einzelne Prüfungsleistungen, aber auch der Hochschulabschluss insgesamt, nachträglich aberkannt werden (vgl. § 33 ASPO).

- **Eigenleistung**

Bewertet werden kann nur die studentische Eigenleistung. Die exzessive Nutzung von KI-Tools kann dazu führen, dass die verbleibende Eigenleistung nicht ausreicht, um die geforderten Kompetenzen erfolgreich zu demonstrieren.

05. WEITERFÜHRENDES MATERIAL

- ❖ Begriffsübersicht: Erläuterung vieler geläufiger Begriffe und Konzepte im Kontext künstlicher Intelligenz: <https://www.ki.nrw/ki-schluesselbegriffe>
- ❖ Toolübersicht: Übersicht über KI-Tools, die das wissenschaftliche Arbeiten unterstützen können: <https://www.vkkiwa.de/ki-ressourcen/ki-tools>
<https://theresanaiforthat.com/>
- ❖ Toolfinder: KI-Tool, mit dem sich andere KI-Tools finden lassen: <https://ki-suche.io/>
- ❖ Übersicht über Chancen und Risiken von Künstlicher Intelligenz: <https://www.europarl.europa.eu/topics/de/article/20200918STO87404/kunstliche-intelligenz-chancen-und-risiken>
- ❖ Material zu existenziellen Risiken von Künstlicher Intelligenz: <https://futureoflife.org/resource/introductory-resources-on-ai-risks/>
- ❖ Übersicht über Anwendungsgebiete von Künstlicher Intelligenz: <https://research.aimultiple.com/generative-ai-applications/>
- ❖ Handreichung ChatGPT der HFT Stuttgart: <https://wiki.hft-stuttgart.de/pages/viewpage.action?spaceKey=NNHS&title=Handreichung+ChatGPT#>

06. INKRAFTTRETEN

Diese Leitlinie tritt am 11.11.2025 nach der Beschlussfassung durch das Rektorat in Kraft. Sie gilt ab diesem Tag.

Stuttgart, 11.11.2025

Prof. Dr. Katja Rade

Rektorin

Bekanntmachungsnachweis
Aushang am:
Abgenommen am:

Beurkundung: