

Ethikrichtlinien für den Umgang mit Künstlicher Intelligenz (KI) am IFP Basel

1.0 Einleitung & Zielsetzung	2
1.1 Rechtliche und ethische Grundlagen.....	2
2.0 Grundprinzipien der KI-Nutzung am IFP Basel	3
2.1 Leitlinien zur verantwortungsvollen Nutzung von KI.....	3
2.2 Verpflichtungen der Organisation und der KI-Nutzer:innen.....	4
2.3 Regelmässige Überprüfung der KI-Prinzipien.....	4
3.0 Anwendungsbereiche der KI am IFP Basel	5
3.1 KI in der Aufbereitung von Lehrmaterialien.....	5
Ethik-Richtlinien für den Einsatz von KI in der Aufbereitung von Lehrmaterialien.....	5
3.2 KI zur Unterstützung bei Leistungsbewertungen.....	7
Ethik-Richtlinien für den Einsatz von KI zur Unterstützung bei Leistungsbewertungen.	8
3.3 KI zur Recherche & strategischen Entscheidungsfindung.....	10
Ethik-Richtlinien für den Einsatz von KI zur Recherche & strategischen Entscheidungsfindung.....	11
4.0 Einrichtung des KI-Rates am IFP Basel	13
4.1 Struktur des KI-Rates.....	13
4.2 Aufgaben und Verantwortlichkeiten des KI-Rates.....	13
4.3 Prozess der KI-Tool-Zulassung.....	14
4.4 Berichtspflicht & Transparenz.....	14
5.0 Transparenz- und Fairnesspflichten	15
5.1 Verpflichtungen zur Transparenz und Fairness.....	15
5.2 Erweiterte Transparenzmassnahmen.....	17
5.3 Massnahmen zur Fairnesssicherung.....	17
6.0 Datenschutz & Sicherheit	18
6.1 Datenschutzrichtlinien für KI-Anwendungen.....	18
6.2 Sicherheitsmassnahmen für KI-Systeme.....	19
6.3 Massnahmen zur Datenschuttschulung & Sensibilisierung.....	20
6.4 Massnahmen zur Einhaltung und Kontrolle.....	20
7.0 Schulungs- & Kontrollmechanismen	21
7.1 Schulungsmassnahmen für KI-Nutzer:innen.....	21
7.2 Kontrollmechanismen & Audits.....	23
7.3 Feedbackmechanismen & kontinuierliche Verbesserung.....	23
7.4 Integration der KI-Thematik in den internen Auditplan.....	24

1.0 Einleitung & Zielsetzung

Das Institut für Führung und Psychologie (IFP Basel) verpflichtet sich zur verantwortungsvollen und ethischen Nutzung von Künstlicher Intelligenz (KI) in allen Bereichen, in denen KI-Systeme eingesetzt werden. Diese Ethikrichtlinien dienen als Ergänzung zum bestehenden Code of Ethics des IFP Basel und sollen sicherstellen, dass der Einsatz von KI menschenzentriert, fair, transparent und nachhaltig erfolgt.

1.1 Rechtliche und ethische Grundlagen

Diese Richtlinien orientieren sich an:

- a) EU AI Act (Regulierung für Künstliche Intelligenz in der EU)
- b) ISO/IEC 42001:2023 (KI-Managementsysteme)
- c) DSGVO & DSG (Datenschutzanforderungen)
- d) Den bestehenden Ethikrichtlinien des IFP Basel
- e) Ethische Leitlinien der UNESCO zur Künstlichen Intelligenz
- f) Empfehlungen des Schweizerischen Datenschutzbeauftragten für den KI-Einsatz

Diese Richtlinie soll sicherstellen, dass KI nicht nur als technische Unterstützung gesehen wird, sondern in einem menschlichen, sozialen und ethischen Kontext eingesetzt wird.

2.0 Grundprinzipien der KI-Nutzung am IFP Basel

Unsere KI-Systeme dienen der Unterstützung und nicht dem Ersatz menschlicher Entscheidungsfindung.

2.1 Leitlinien zur verantwortungsvollen Nutzung von KI

Grundprinzip	Beschreibung
Förderung der menschlichen Kreativität und sozialen Interaktion	KI soll Lernprozesse unterstützen, indem sie Studierenden und Lehrkräften hilft, effizienter zu arbeiten und sich auf kreative sowie analytische Aufgaben zu konzentrieren. Sie darf jedoch keine schöpferische oder interaktive Kompetenz ersetzen.
Transparenz und Fairness	Alle KI-gestützten Prozesse müssen nachvollziehbar und für Betroffene verständlich sein. Nutzer:innen müssen informiert werden, wenn KI-Entscheidungen eine Rolle spielen.
Diskriminierungsfreiheit	KI-Systeme dürfen keine direkten oder indirekten Benachteiligungen aufgrund von Geschlecht, Herkunft, Alter, Behinderung oder anderen geschützten Merkmalen erzeugen. Dies erfordert eine regelmässige Überprüfung der eingesetzten Algorithmen.
Keine automatisierten, endgültigen Bewertungen oder Bewerbungsentscheidungen	KI kann Bewertungen unterstützen, aber die finale Entscheidung über eine Prüfungsleistung, einen Studienplatz oder eine Bewerberauswahl liegt immer bei einem Menschen.
Datenschutz und Wahrung der Privatsphäre	KI-Systeme müssen die Datenschutzbestimmungen der DSGVO und DSG strikt einhalten. Dies umfasst die Minimierung personenbezogener Datenerhebung, die Verschlüsselung sensibler Daten und die Gewährleistung des Rechts auf Datenlöschung.
Menschliche Kontrolle über KI-Entscheidungen	KI darf nur eine beratende Funktion haben und niemals die alleinige Entscheidungsinstanz sein. Kritische Bewertungen müssen stets durch menschliche Fachkräfte geprüft werden.

Ethische Reflexion und verantwortungsbewusste Entscheidungsunterstützung	KI soll Daten analysieren und Erkenntnisse liefern, darf jedoch nicht eigenständig Bewertungen oder ethische Urteile fällen. Dies bleibt der menschlichen Intuition und Expertise überlassen.
Unterstützung des Lernverhaltens ohne Manipulation	KI kann personalisierte Lernempfehlungen geben, jedoch dürfen keine Mechanismen eingesetzt werden, die bewusst psychologische Manipulation fördern oder das Lernverhalten in eine bestimmte Richtung lenken.

2.2 Verpflichtungen der Organisation und der KI-Nutzer:innen

Verantwortlich	Pflichten
Führungsebene	Sicherstellen, dass alle KI-Prozesse ethischen Standards und gesetzlichen Anforderungen entsprechen. Förderung einer offenen KI-Kultur mit klaren Kommunikationslinien zu Risiken und Nutzen.
Lehrkräfte & Prüfer:innen	Sensibler und informierter Umgang mit KI-gestützten Bewertungssystemen, Sicherstellung der finalen menschlichen Kontrolle. Regelmässige Fortbildung zur KI-Nutzung.
Studierende & Mitarbeitende	Bewusstsein für den Einsatz von KI im akademischen Umfeld. Möglichkeit zur Hinterfragung und kritischen Auseinandersetzung mit KI-gestützten Ergebnissen.
IT- & Datenverantwortliche	Implementierung sicherer, nachvollziehbarer und fairer KI-Systeme. Regelmässige technische Überprüfung der Algorithmen auf Bias und Datensicherheit.

2.3 Regelmässige Überprüfung der KI-Prinzipien

- a) Jährliche Evaluation der eingesetzten KI-Systeme durch interne Audits und externe Prüfinstanzen.
- b) Feedbackmechanismen für Studierende, Lehrkräfte und Verwaltungspersonal zur kontinuierlichen Verbesserung der KI-Nutzung.
- c) Anpassung der Richtlinien auf Basis neuer regulatorischer Vorgaben,

technologischer Entwicklungen und ethischer Standards.

3.0 Anwendungsbereiche der KI am IFP Basel

Die Nutzung von KI beschränkt sich auf drei zentrale Bereiche:

3.1 KI in der Aufbereitung von Lehrmaterialien

Einsatzbereich	Beschreibung
KI-gestützte Recherche & Literaturanalyse	Unterstützung von Lehrkräften durch effiziente Datenverarbeitung und Literaturrecherche zu aktuellen wissenschaftlichen Themen.
Generierung von Podcast-Inhalten und digitalen Lernressourcen	KI kann Inhalte strukturieren und Zusammenfassungen erstellen, die in Podcasts oder digitale Lernplattformen integriert werden.
Erstellung und Optimierung von Lernmaterialien	KI kann Inhalte für unterschiedliche Lernstile aufbereiten und personalisierte Selbststudienmaterialien erstellen.
Erstellung wissenschaftlicher Zusammenfassungen und visuelle Lernmaterialien	Unterstützung bei der Strukturierung und Visualisierung komplexer Inhalte, z. B. durch Diagramme oder interaktive Lernelemente.

Ethik-Richtlinien für den Einsatz von KI in der Aufbereitung von Lehrmaterialien

Um sicherzustellen, dass die Nutzung von KI zur Unterstützung der Lehrmaterialien am IFP Basel **ethisch vertretbar, transparent und fachlich korrekt** erfolgt, gelten die folgenden verbindlichen Richtlinien:

a) **Menschliche Endkontrolle durch qualifizierte Fachkräfte:**

- Alle durch KI erstellten oder aufbereiteten Lehrmaterialien müssen vor ihrer Verwendung von **qualifizierten Fachkräften geprüft und freigegeben werden**.
- Diese Prüfung erfolgt gemäss den Vorgaben der **Lehrpersonenverordnung** und stellt sicher, dass die Inhalte **fachlich korrekt, didaktisch geeignet und frei von Verzerrungen oder Fehlinterpretationen** sind.

b) Kennzeichnung von KI-generierten Inhalten zur Wahrung der Transparenz:

- Jede durch KI generierte oder unterstützte Lehrressource muss deutlich **gekennzeichnet** werden, damit Studierende und Lehrkräfte jederzeit erkennen können, ob und in welchem Umfang KI an der Erstellung beteiligt war.
- Dies kann durch einen Hinweis wie *„Dieser Inhalt wurde mit KI-Unterstützung durch ausgewiesene akademischen Fachpersonen erstellt und durch Fachkräfte überprüft“* erfolgen.
- Die Kennzeichnungspflicht dient der **Transparenz** und verhindert mögliche **Fehlinterpretationen oder Täuschungen**.

c) Menschliche Validierung jeder KI-gestützten Materialerstellung, insbesondere bei prüfungsrelevanten Inhalten:

- Lehrkräfte sind verpflichtet, **alle KI-gestützten Inhalte, die in Prüfungen oder benoteten Lernmaterialien verwendet werden**, kritisch zu prüfen.
- Eine **automatische Erstellung von Prüfungsmaterialien durch KI ist nicht zulässig**, es sei denn, sie wurde durch eine qualifizierte Lehrkraft **vollständig überprüft und freigegeben**.
- Studierende müssen sich darauf verlassen können, dass alle prüfungsrelevanten Materialien **methodisch und inhaltlich einwandfrei** sind.

d) Verbot der Manipulation oder Verzerrung von Inhalten durch KI-Modelle:

- KI-Systeme dürfen nicht so eingesetzt werden, dass sie Inhalte **absichtlich verändern, manipulieren oder verzerren**, um bestimmte Meinungen oder Perspektiven zu bevorzugen.
- Alle generierten Texte oder Visualisierungen müssen **den wissenschaftlichen und didaktischen Standards** des IFP Basel entsprechen und dürfen **keine Voreingenommenheit oder Falschinformationen enthalten**.

- Falls eine KI **unbeabsichtigt falsche oder verzerrte Inhalte generiert**, müssen diese umgehend korrigiert oder entfernt werden.

e) KI darf nicht zur Generierung von Lehrmeinungen oder wissenschaftlichen Theorien verwendet werden, sondern nur unterstützend wirken:

- Die KI darf **keine eigenen wissenschaftlichen Theorien, Lehrmeinungen oder Forschungsergebnisse erstellen oder verbreiten**.
- Ihr Einsatz ist **ausschliesslich unterstützender Natur**, beispielsweise zur **Strukturierung, Analyse oder Zusammenfassung bestehender wissenschaftlicher Inhalte**.
- Wissenschaftliche Aussagen und Interpretationen dürfen nur von **menschlichen Expert:innen vorgenommen** werden.

Diese Richtlinien stellen sicher, dass KI am IFP Basel **ethisch korrekt, fachlich zuverlässig und transparent** eingesetzt wird. Gleichzeitig wird gewährleistet, dass **menschliche Fachkräfte stets die Verantwortung für die Qualität und Richtigkeit der Lehrmaterialien tragen**.

3.2 KI zur Unterstützung bei Leistungsbewertungen

Einsatzbereich	Beschreibung
Objektivierung von Prüfungsbewertungen	KI kann als Unterstützung bei der Bewertung von Prüfungsleistungen dienen, indem sie Muster und Abweichungen analysiert.
Überprüfung von Prüfungsleistungen mit Analysetools	KI kann Plagiate, fehlerhafte Zitierungen oder sprachliche Muster erkennen, die auf Unsicherheiten hinweisen.
Unterstützung bei plagiatsbezogenen Prüfungen	KI-Systeme können frühzeitig potenzielle Plagiate in schriftlichen Arbeiten aufdecken.

Ethik-Richtlinien für den Einsatz von KI zur Unterstützung bei Leistungsbewertungen

Um sicherzustellen, dass der Einsatz von KI in Prüfungs- und Bewertungsprozessen am IFP Basel **fair, transparent und unterstützend** erfolgt, gelten die folgenden verbindlichen Richtlinien:

a) **KI dient ausschliesslich der Unterstützung und ersetzt keine menschlichen Bewertungen:**

- KI kann zur **Unterstützung von Prüfer:innen** eingesetzt werden, beispielsweise zur **Erkennung von Mustern, zur Analyse von Texten oder zur Überprüfung auf Plagiate**.
- Die endgültige Bewertung einer Prüfungsleistung oder einer wissenschaftlichen Arbeit muss jedoch **immer von einer menschlichen Fachkraft vorgenommen werden**.
- KI darf keine **alleinige Bewertungsinstanz sein oder automatische Benotungen ohne menschliche Kontrolle durchführen**.

b) **Verpflichtung der Prüfer:innen, KI-Ergebnisse zu hinterfragen und zu validieren:**

- Prüfer:innen sind dazu verpflichtet, **die von KI bereitgestellten Analysen, Vorschläge oder Bewertungen kritisch zu hinterfragen und zu überprüfen**.
- Falls eine KI eine **Voreingenommenheit oder systematische Abweichungen** zeigt, muss dies sofort erkannt und korrigiert werden.
- Es müssen **klare Kriterien und Validierungsprozesse** definiert werden, um sicherzustellen, dass die Nutzung der KI die Qualität der Bewertung nicht beeinträchtigt.

c) **Studierende müssen über den Einsatz von KI in Prüfungsprozessen informiert werden:**

- Alle Studierenden müssen **im Vorfeld über die Verwendung von KI in Prüfungs- und Bewertungsprozessen informiert werden**.
- Die Art und Weise, wie KI zum Einsatz kommt, sowie deren **Funktion und Einschränkungen**, müssen **transparent kommuniziert werden**.

- Studierende haben das **Recht, eine ausschliesslich menschliche Bewertung ihrer Prüfungsleistung zu verlangen**, falls sie Zweifel an der KI-gestützten Unterstützung haben.

d) Jährliche Prüfung der KI-Modelle auf Bias, Fairness und Objektivität:

- Alle KI-Systeme, die zur Unterstützung von Prüfungen und Bewertungen eingesetzt werden, müssen **einmal jährlich einer umfassenden Überprüfung** unterzogen werden.
- Der **KI-Rat** ist dafür verantwortlich, zu evaluieren, ob die verwendeten KI-Modelle **unbeabsichtigte Diskriminierungen (Bias), systematische Fehler oder Unfairness aufweisen**.
- Falls festgestellt wird, dass eine KI **nicht fair oder objektiv bewertet**, muss sie sofort aus dem Prüfungsprozess entfernt oder durch ein verbessertes Modell ersetzt werden.

e) KI darf keine emotionale oder verhaltensbasierte Analyse der Studierenden durchführen:

- KI-Systeme dürfen **keine psychologischen Profile von Studierenden erstellen oder deren emotionale oder soziale Verhaltensweisen interpretieren**.
- Es ist verboten, KI für die **Analyse von Mimik, Gestik, Sprachmustern oder anderen verhaltensbasierten Indikatoren** im Prüfungsumfeld einzusetzen.
- Prüfungen sollen **auf Wissen und Kompetenz basieren**, nicht auf emotionalen oder persönlichen Merkmalen, die durch KI interpretiert werden könnten.

f) Kein Einsatz von KI zur Vorhersage des akademischen Erfolgs oder der persönlichen Entwicklung:

- KI darf nicht verwendet werden, um auf Basis von bisherigen Leistungen, persönlichen Daten oder Verhaltensmustern **Vorhersagen über den akademischen Erfolg oder die persönliche Entwicklung eines Studierenden zu treffen**.
- Entscheidungen über die **Studienlaufbahn, Förderung oder individuelle Unterstützungsmassnahmen** dürfen nicht durch **KI-gestützte Prognosen** getroffen werden, sondern müssen immer auf **menschlicher Einschätzung und Beratung** beruhen.

- Der Einsatz von KI für solche Zwecke birgt **erhebliche ethische Risiken und kann unbeabsichtigte Diskriminierungen oder Fehleinschätzungen** zur Folge haben.

Diese Richtlinien gewährleisten, dass **KI als unterstützendes Werkzeug, jedoch niemals als Ersatz für menschliche Expertise** in Prüfungsprozessen eingesetzt wird. **Die finale Verantwortung für Bewertungen liegt immer bei menschlichen Fachkräften**, um **Fairness, Transparenz und eine individuelle Betrachtung jeder Prüfungsleistung** zu sichern.

3.3 KI zur Recherche & strategischen Entscheidungsfindung

Einsatzbereich	Beschreibung
Datenanalyse & Marktanalysen zur Strategieentwicklung	KI kann helfen, Markttrends zu analysieren, Prognosen zu erstellen und strategische Entscheidungen mit datenbasierten Erkenntnissen zu unterstützen.
Textgenerierung für Marketing- und Werbepublikationen	KI kann Vorschläge für Werbetexte, Kampagneninhalte oder Social-Media-Beiträge generieren, die von Marketingteams überprüft und angepasst werden.
Optimierung interner Verwaltungsprozesse	KI kann Routineaufgaben in der Verwaltung automatisieren, wie Dokumentenmanagement oder Terminplanung.
Erstellung automatisierter Berichte und Dokumentationen	KI kann unterstützend Berichte aufbereiten, die anschliessend von Fachkräften überprüft werden.

Ethik-Richtlinien für den Einsatz von KI zur Recherche & strategischen Entscheidungsfindung

Um sicherzustellen, dass KI am IFP Basel in den Bereichen **Recherche, Marktanalyse und strategische Entscheidungsfindung** verantwortungsvoll und ethisch korrekt eingesetzt wird, gelten die folgenden verbindlichen Richtlinien:

a) KI-generierte Texte dürfen nicht ohne menschliche Endkontrolle veröffentlicht werden:

- Texte, Berichte, Marktanalysen oder strategische Empfehlungen, die unter Verwendung von KI erstellt wurden, müssen vor der Veröffentlichung **durch menschliche Fachkräfte überprüft und freigegeben werden.**
- Dies gilt insbesondere für **wissenschaftliche Publikationen, offizielle Dokumente, Lehrmaterialien und Marketingtexte**, um sicherzustellen, dass keine **fehlerhaften, einseitigen oder manipulativen Inhalte verbreitet werden.**
- Vor der Veröffentlichung muss ein **klarer Validierungsprozess** durch Fachkräfte definiert sein, der sicherstellt, dass **alle Aussagen den geltenden wissenschaftlichen und ethischen Standards entsprechen.**

b) Alle KI-gestützten Analysen müssen durch menschliche Experten validiert werden:

- KI kann als Unterstützung zur **Datenanalyse, Marktprognose oder strategischen Entscheidungsfindung** genutzt werden, aber die finalen Schlüsse dürfen **nur durch menschliche Experten gezogen werden.**
- Jede KI-gestützte Analyse muss von einer **fachlich qualifizierten Person überprüft und kritisch hinterfragt** werden, bevor sie für Management- oder Strategieentscheidungen genutzt wird.
- Entscheidungen, die auf **KI-generierten Analysen basieren, müssen dokumentiert** und mit alternativen Datenquellen abgeglichen werden, um Verzerrungen oder Fehleinschätzungen zu vermeiden.

c) Einhaltung der Datenschutzbestimmungen gemäss DSGVO & DSG:

- Alle KI-gestützten Recherchen und Analysen müssen strikt **den Vorgaben der Datenschutz-Grundverordnung (DSGVO) und des Schweizer Datenschutzgesetzes (DSG)** entsprechen.
- **Es dürfen keine personenbezogenen Daten ohne ausdrückliche Einwilligung verarbeitet, gespeichert oder analysiert werden.**
- Falls KI-gestützte Systeme auf **personenbezogene oder sensible Daten zugreifen**, muss eine **Datenschutz-Folgenabschätzung (DPIA) durchgeführt und dokumentiert werden.**
- Nutzer:innen müssen jederzeit das **Recht auf Einsicht, Berichtigung und**

Löschung ihrer Daten haben, wenn diese in KI-gestützten Analysen verwendet wurden.

d) KI darf keine personalisierten Profile von Studierenden oder Mitarbeitenden erstellen:

- Es ist verboten, durch KI gestützte Algorithmen zur **Erstellung individueller Profile von Studierenden oder Mitarbeitenden** zu nutzen, die auf deren Verhalten, akademischen Leistungen oder persönlichen Präferenzen basieren.
- Solche Systeme könnten zu **verdeckten Bewertungen, Diskriminierung oder unbeabsichtigten Fehlinterpretationen** führen und sind daher am IFP Basel nicht zulässig.
- Falls eine **individuelle Empfehlung für Studierende oder Mitarbeitende** ausgesprochen wird (z. B. für Studiengänge, Fortbildungen oder Arbeitsbereiche), muss diese **auf expliziter menschlicher Einschätzung basieren** und darf nicht durch KI-generierte Profile ersetzt werden.

e) Verbot der automatisierten Meinungsbildung durch KI-basierte Analysen:

- KI darf **keine automatisierte Meinung oder Wertung über eine Person, ein Forschungsthema oder eine strategische Richtung erstellen**, die als Grundlage für offizielle Entscheidungen dient.
- Alle durch KI generierten Inhalte und Analysen dürfen **ausschliesslich als Unterstützung** für Entscheidungsprozesse genutzt werden, müssen aber **durch menschliche Expertise ergänzt und bewertet werden**.
- Dies stellt sicher, dass **kein KI-System aktiv Meinungen oder Richtlinien des IFP Basel beeinflussen kann**.

f) Menschliche Entscheidungsträger müssen in alle kritischen Prozesse eingebunden sein:

- **KI darf niemals als alleinige Entscheidungsinstanz fungieren**. Alle kritischen Entscheidungen – insbesondere in den Bereichen **Personalmanagement, strategische Ausrichtung, Forschung, Prüfungswesen und ethische Richtlinien** – müssen **durch menschliche Verantwortliche geprüft und getroffen werden**.
- Der **KI-Rat trägt die Verantwortung**, sicherzustellen, dass alle KI-gestützten

Entscheidungsprozesse **dokumentiert und transparent gemacht werden.**

- Falls KI ein Fehlverhalten oder systematische Verzerrungen aufweist, ist eine **sofortige Überprüfung und gegebenenfalls die Deaktivierung der betreffenden KI-Systeme erforderlich.**

4.0 Einrichtung des KI-Rates am IFP Basel

Um den ethischen, sicheren und regelkonformen Einsatz von Künstlicher Intelligenz (KI) am IFP Basel sicherzustellen, wird ein KI-Rat als überwachendes Gremium eingerichtet.

4.1 Struktur des KI-Rates

- a) Der KI-Rat besteht aus 3 Mitgliedern und einem permanenten Vorsitzenden.
- b) Das Gremium ist direkt der Ombudsstelle (OO - Ombudsman Office) des IFP Basel unterstellt.
- c) Die Mitglieder setzen sich aus interdisziplinären Experten aus den Bereichen Qualitätsmanagement, IT-Sicherheit, Datenschutz und Didaktik zusammen.
- d) Der Vorsitzende koordiniert die Sitzungen, Auditierungen und Berichte zur KI-Nutzung.

4.2 Aufgaben und Verantwortlichkeiten des KI-Rates

Aufgabe	Beschreibung
Auswahl zulässiger KI-Tools	Der KI-Rat prüft und genehmigt KI-Anwendungen, die am IFP Basel eingesetzt werden dürfen.
Regelmässige Prüfung & Auditierung	Jährliche Bewertung aller genehmigten KI-Tools hinsichtlich Fairness, Datenschutz, Transparenz und ethischer Verträglichkeit.
Schulung der Nutzer:innen	Organisation und Durchführung von Fortbildungen für alle Mitarbeitenden zu den zugelassenen KI-Tools und deren korrekter Anwendung.
Führung einer KI-Tool-Liste	Verwaltung einer zentralen Liste aller zugelassenen und geprüften

	KI-Anwendungen am IFP Basel.
Risikobewertung der KI-Tools	Regelmässige Durchführung von Risikobewertungen und Analysen, um mögliche Gefahren und Missbrauchspotenziale frühzeitig zu erkennen.
Festlegung der konkreten Anwendung	Definition und Dokumentation, in welchen Bereichen und unter welchen Bedingungen die KI-Tools eingesetzt werden dürfen.
Überprüfung der Einhaltung der Ethikrichtlinien	Sicherstellung, dass alle eingesetzten KI-Tools den festgelegten ethischen Richtlinien, Datenschutzbestimmungen und Fairnesskriterien entsprechen.

4.3 Prozess der KI-Tool-Zulassung

- a) Antragstellung: Neue KI-Tools werden durch Lehrkräfte oder Verwaltungsmitarbeiter vorgeschlagen.
- b) Prüfung durch den KI-Rat: Analyse des Tools hinsichtlich Datenschutz, Ethik, Bias-Risiken und Transparenz.
- c) Freigabe oder Ablehnung: Der KI-Rat entscheidet über die Zulassung oder Ablehnung eines Tools.
- d) Schulungen & Dokumentation: Schulung der Nutzer und Aufnahme in die offizielle KI-Tool-Liste des IFP Basel.
- e) Regelmässige Überprüfung: Jährliche Auditierung aller zugelassenen KI-Tools durch den KI-Rat.

4.4 Berichtspflicht & Transparenz

- a) Der KI-Rat erstellt jährlich einen öffentlichen Bericht über die Nutzung, Überprüfung und Evaluierung der KI-Tools am IFP Basel.
- b) Alle Studierenden und Mitarbeitenden haben das Recht auf Einsicht in die genehmigte KI-Tool-Liste.
- c) Die Ergebnisse der KI-Auditierungen werden in das interne Qualitätsmanagement integriert und in den jährlichen Management Review aufgenommen.

5.0 Transparenz- und Fairnesspflichten

Die Transparenz- und Fairnesspflichten stellen sicher, dass der Einsatz von KI-Systemen am IFP Basel nachvollziehbar, diskriminierungsfrei und ethisch verantwortungsvoll erfolgt. Es werden Massnahmen definiert, um menschliche Kontrolle, regelmässige Überprüfung und den Schutz vor Diskriminierung zu gewährleisten.

5.1 Verpflichtungen zur Transparenz und Fairness

Verpflichtung	Beschreibung	Konkrete Massnahmen
Offenlegung der KI-Nutzung	Jede Nutzung von KI muss für alle Betroffenen transparent gemacht werden.	<ul style="list-style-type: none"> - Alle KI-gestützten Inhalte (z. B. Prüfungsbewertungen, Lehrmaterialien, administrative Texte) müssen gekennzeichnet sein. - Standardhinweis: „Dieser Text wurde mit KI-Unterstützung erstellt.“ - Einführung eines KI-Nutzungsregisters, das dokumentiert, wann und wo KI eingesetzt wurde.
Diskriminierungsfreiheit	KI-Systeme dürfen keine Benachteiligungen aufgrund von Geschlecht, Herkunft, Alter oder anderen persönlichen Merkmalen erzeugen.	<ul style="list-style-type: none"> - Regelmässige Bias-Überprüfung durch den KI-Rat. - Implementierung von Anti-Diskriminierungsfiltren in KI-gestützten Bewertungssystemen. - Dokumentierte Massnahmen zur Bias-Reduktion (z. B. diverse Trainingsdatensätze, Tests auf Fairness).

Menschliche Kontrolle	KI darf keine endgültigen Entscheidungen treffen; es muss immer eine menschliche Überprüfung stattfinden.	<ul style="list-style-type: none"> - Einführung eines Vier-Augen-Prinzips bei KI-gestützten Prüfungsbewertungen. - Schulung von Prüfer:innen, um KI-Entscheidungen kritisch zu hinterfragen. - Studierende & Mitarbeitende haben das Recht, eine menschliche Entscheidung anzufordern, falls sie mit einer KI-gestützten Einschätzung nicht einverstanden sind.
Jährliche Überprüfung	Die Qualität und Fairness der eingesetzten KI-Systeme wird regelmässig auditiert.	<ul style="list-style-type: none"> - Jährliches KI-Tool-Audit durch den KI-Rat. - Überprüfung von Leistungsmetriken, Nutzerfeedback und Bias-Analysen. - Falls sich Unregelmässigkeiten zeigen, wird die Nutzung des betreffenden KI-Tools gestoppt oder angepasst.
Zugang zu Trainingsdaten	Nutzer:innen müssen wissen, auf welchen Daten KI-Modelle basieren.	<ul style="list-style-type: none"> - Veröffentlichung einer Dokumentation über verwendete Trainingsdaten. - Möglichkeit für Studierende & Mitarbeitende, eine Erklärung zur Funktionsweise von KI-gestützten Prozessen anzufordern. - Falls ein KI-System auf sensiblen oder personenbezogenen Daten basiert, müssen datenschutzrechtliche Massnahmen getroffen werden.

5.2 Erweiterte Transparenzmassnahmen

Um sicherzustellen, dass KI-Prozesse klar verständlich und offen dokumentiert sind, werden folgende Mechanismen eingeführt:

- a) KI-Transparenzbericht: Jährliche Veröffentlichung eines Berichts zur Nutzung von KI am IFP Basel.
- b) KI-Labeling-System: Kennzeichnung aller Inhalte, die mit KI generiert, überarbeitet oder geprüft wurden.
- c) Beschwerde- und Einspruchsmöglichkeit: Nutzer:innen können eine menschliche Überprüfung beantragen, wenn sie eine KI-basierte Entscheidung als fehlerhaft oder unfair empfinden.

5.3 Massnahmen zur Fairnesssicherung

KI-Systeme werden regelmässig auf ethische & faire Entscheidungsfindung geprüft. Der KI-Rat führt regelmässige Tests durch, um sicherzustellen, dass:

- a) Keine unbeabsichtigten Diskriminierungen oder Verzerrungen entstehen.
- b) Menschliche Kontrolle über KI-gestützte Prozesse gewährleistet ist.
- c) Alle Nutzer:innen über den KI-Einsatz informiert und geschult sind.

Falls ein KI-Tool diese Standards nicht erfüllt, wird es ausgesetzt, überarbeitet oder ganz abgeschafft.

6.0 Datenschutz & Sicherheit

Der Datenschutz und die Sicherheit von KI-Systemen sind essenziell für den vertrauenswürdigen Einsatz am IFP Basel. KI-Anwendungen müssen sicherstellen, dass alle personenbezogenen Daten geschützt sind, keine unrechtmässige Verarbeitung erfolgt und alle Sicherheitsstandards eingehalten werden.

6.1 Datenschutzrichtlinien für KI-Anwendungen

Datenschutzrichtlinie	Beschreibung	Konkrete Massnahmen
Einhaltung von DSGVO & DSG	Alle KI-Systeme müssen den Datenschutzanforderungen der EU-Datenschutzgrundverordnung (DSGVO) und des Schweizer Datenschutzgesetzes (DSG) entsprechen.	<ul style="list-style-type: none"> - Verarbeitung nur mit Rechtsgrundlage (Einwilligung oder gesetzliche Notwendigkeit). - Datenschutz-Folgenabschätzung (DPIA) für KI-gestützte Prozesse. - Dokumentation der KI-Datenverarbeitung gemäss DSGVO Art. 30.
Keine Verarbeitung ohne Zustimmung	Personenbezogene Daten dürfen nur verarbeitet werden, wenn eine explizite Einwilligung der betroffenen Person vorliegt.	<ul style="list-style-type: none"> - Opt-in Verfahren für alle personenbezogenen Daten in KI-gestützten Prozessen. - Nutzerfreundliche Einwilligungsoptionen, die jederzeit widerrufen werden können. - Datenschutzfreundliche Voreinstellungen („Privacy by Default“) für alle KI-Anwendungen.
Recht auf Löschung & Korrektur	Nutzer:innen haben das Recht, ihre Daten löschen oder korrigieren zu lassen, wenn diese fehlerhaft oder nicht mehr notwendig sind.	<ul style="list-style-type: none"> - Automatisierte Prozesse für Datenlöschung nach definierten Fristen. - Klar definierte Ansprechpartner für Datenschutzanfragen. - Jährliche Kontrolle der gespeicherten Daten durch den KI-Rat.
Minimierung & Anonymisierung	Es dürfen nur die absolut notwendigen Daten verarbeitet werden. Personenbezogene Daten sollten wenn möglich anonymisiert oder pseudonymisiert werden.	<ul style="list-style-type: none"> - KI-gestützte Prozesse müssen ohne Identitätsbezug funktionieren, falls dies nicht erforderlich ist. - Automatische Anonymisierungsskripte für Trainingsdaten. - Pseudonymisierung sensibler Daten in KI-Systemen zur Risikominimierung.

Transparenz über Datennutzung	Nutzer:innen müssen genau wissen, welche ihrer Daten verarbeitet werden.	<ul style="list-style-type: none"> - Erstellung eines „KI-Datenverarbeitungsverzeichnisses“, das alle gespeicherten und verarbeiteten Daten aufführt. - Einfache, verständliche Datenschutzerklärungen zu KI-Prozessen. - Regelmässige Audits zur Kontrolle der Datenverarbeitung.
-------------------------------	--	---

6.2 Sicherheitsmassnahmen für KI-Systeme

Sicherheitsmassnahme	Beschreibung	Konkrete Umsetzung
ISO 27001-Zertifizierung	Alle Systeme, die personenbezogene Daten speichern oder verarbeiten, müssen auf ISO 27001-zertifizierten Servern laufen.	<ul style="list-style-type: none"> - Speicherung von Daten ausschliesslich auf sicheren Rechenzentren mit Zugriffsprotokollen & Verschlüsselung. - Jährliche Überprüfung der Server-Sicherheitszertifizierung durch IT-Sicherheitsbeauftragte.
Regelmässige Sicherheitsaudits	KI-Systeme müssen regelmässig auf Sicherheitslücken und potenzielle Bedrohungen überprüft werden.	<ul style="list-style-type: none"> - Halbjährliche Sicherheitsaudits durch interne und externe Prüfer. - Penetrationstests für KI-Systeme, um Angriffsvektoren zu identifizieren. - Verpflichtende Reaktionsstrategie auf Sicherheitsvorfälle.
Schutz gegen Cyberangriffe	KI-Systeme sind potenzielle Angriffsziele. Massnahmen zur Erkennung und Abwehr von Cyberangriffen müssen implementiert werden.	<ul style="list-style-type: none"> - Echtzeit-Überwachung und Logging von Zugriffen auf KI-Systeme. - Multi-Faktor-Authentifizierung für alle administrativen Zugänge. - Firewall- und Intrusion-Detection-Systeme für alle KI-gestützten Anwendungen.
Zugriffsmanagement & Rollenvergabe	Es müssen strenge Regeln dafür existieren, wer Zugriff auf KI-Daten und -Systeme hat.	<ul style="list-style-type: none"> - Nur autorisierte Personen haben Zugriff auf sensible KI-Daten. - Rollenbasiertes Zugriffskontrollsystem (RBAC) zur Verhinderung von Datenlecks. - Protokollierung aller Datenzugriffe & Änderungsprozesse.

Sichere Datenübertragung & -speicherung	KI-Systeme dürfen keine ungesicherten Datenübertragungen oder -speicherungen zulassen.	<ul style="list-style-type: none"> - Ende-zu-Ende-Verschlüsselung für alle KI-gestützten Datenübertragungen. - Speicherung nur auf sicheren, DSGVO-konformen Cloud-Diensten. - Regelmässige Backups mit sicheren Wiederherstellungsmechanismen.
---	--	--

6.3 Massnahmen zur Datenschutzschulung & Sensibilisierung

Um sicherzustellen, dass alle Mitarbeitenden den richtigen Umgang mit KI und Datenschutz kennen, werden folgende Massnahmen eingeführt:

- a) Pflichtschulungen für alle Mitarbeitenden & Studierenden zur sicheren Nutzung von KI-Systemen.
- b) Jährliche Datenschutz-Workshops für Lehrkräfte, Prüfer:innen & Verwaltungsangestellte.
- c) Interne Datenschutz-Checklisten für alle KI-Projekte.
- d) Regelmässige Sicherheitsübungen & Notfalltests, um Datenschutzverletzungen vorzubeugen.

6.4 Massnahmen zur Einhaltung und Kontrolle

Massnahme	Beschreibung	Verantwortlich
Jährliche Datenschutzprüfung	Überprüfung aller KI-Systeme auf Datenschutzkonformität.	KI-Rat & Datenschutzbeauftragter
Interne & externe Audits	Unabhängige Überprüfung der Datenschutzmassnahmen durch externe Prüfer.	Externe Datenschutz- & Sicherheitsfachkräfte
Beschwerdemechanismus	Einrichtung eines anonymen Meldekanals für Datenschutzbedenken.	Ombudsstelle (OO) des IFP Basel
Dokumentation aller Datenschutzvorfälle	Pflicht zur Meldung und Untersuchung aller sicherheitsrelevanten Vorfälle.	IT-Verantwortlicher & Datenschutzbeauftragter

Regelmässige Updates & Anpassung	Laufende Aktualisierung der Datenschutzmassnahmen basierend auf neuen Bedrohungen oder gesetzlichen Anforderungen.	IT-Verantwortlicher & Datenschutzbeauftragter
----------------------------------	--	---

7.0 Schulungs- & Kontrollmechanismen

Um die sichere, ethische und regelkonforme Nutzung von KI am IFP Basel sicherzustellen, sind Schulungs- und Kontrollmechanismen unerlässlich. Diese Massnahmen gewährleisten, dass alle Mitarbeitenden und Studierenden den Umgang mit KI verstehen, Risiken minimiert und kontinuierliche Verbesserungen umgesetzt werden.

Der KI-Rat trägt die Hauptverantwortung für die Entwicklung, Durchführung und Überwachung dieser Massnahmen.

7.1 Schulungsmassnahmen für KI-Nutzer:innen

Alle Nutzer:innen von KI am IFP Basel müssen geschult werden, um sicherzustellen, dass sie die technischen, ethischen und regulatorischen Anforderungen verstehen.

Zielgruppe	Inhalte der Schulung	Häufigkeit	Verantwortlich
Führungskräfte & KI-Entscheidungsträger:innen	<ul style="list-style-type: none"> - Überblick über den EU AI Act und die ISO 42001 - Verantwortlichkeiten bei der KI-Nutzung - Entscheidungsfindung bei KI-gestützten Prozessen - Risikoanalyse für neue KI-Tools 	1x pro Jahr + bei Einführung neuer KI-Tools	KI-Rat

Lehrkräfte & Prüfer:innen	<ul style="list-style-type: none"> - Einsatz von KI in Lehrmaterialien & Prüfungen - Sicherstellung von Fairness & Transparenz - Menschliche Kontrolle & Validierung von KI-generierten Bewertungen 	1x pro Jahr	KI-Rat & Fachschaftratsrat
---------------------------	--	-------------	----------------------------

Alle Schulungen werden dokumentiert und sind verpflichtend für alle, die mit KI-Systemen arbeiten.

7.2 Kontrollmechanismen & Audits

Regelmässige Audits & Überprüfungen gewährleisten, dass KI-Anwendungen weiterhin den ethischen, regulatorischen und sicherheitstechnischen Standards entsprechen.

Auditmassnahme	Zielsetzung	Frequenz	Verantwortlich
Jährliche KI-Compliance-Audits	Überprüfung, ob alle KI-Prozesse DSGVO, DSG & EU AI Act konform sind	1x pro Jahr	KI-Rat & Datenschutzbeauftragter
Bias- und Fairness-Analysen	Sicherstellen, dass KI-Systeme diskriminierungsfrei arbeiten	Alle 6 Monate	KI-Rat & interne/externe Prüfer:innen
Risikobewertung der KI-Tools	Identifikation potenzieller Gefahren und Sicherheitsrisiken bei bestehenden Tools	1x pro Jahr	KI-Rat & IT-Verantwortlicher

Interne Kontrollen durch den KI-Rat	Überprüfung, ob alle Schulungs- und Transparenzmassnahmen umgesetzt werden	Quartalsweise	KI-Rat
-------------------------------------	--	---------------	--------

Falls ein KI-Tool die Audits nicht besteht, wird seine Nutzung ausgesetzt oder angepasst.

7.3 Feedbackmechanismen & kontinuierliche Verbesserung

Ein strukturiertes Feedbacksystem stellt sicher, dass Nutzer:innen ihre Erfahrungen mit KI-gestützten Systemen melden und die Prozesse kontinuierlich verbessert werden können.

Feedback-Methode	Zielsetzung	Verantwortlich
Anonyme KI-Nutzungsumfragen	Erfassung der Zufriedenheit & Probleme mit KI-gestützten Prozessen	KI-Rat
Meldeformular für fehlerhafte oder unfair bewertete KI-Ergebnisse	Möglichkeit für Studierende & Mitarbeitende, Fehlentscheidungen oder Bias zu melden	KI-Rat & Ombudsstelle (OO)
Jährlicher KI-Nutzungsbericht	Veröffentlichung eines Berichts über Einsätze, Audits & Verbesserungen	KI-Rat & Qualitätsmanagement
Feedback-Runde mit Studierenden & Dozierenden	Direkte Einbindung der Nutzer:innen in die Weiterentwicklung der KI-Strategie	KI-Rat & Fachschaftratsrat

Diese Feedbackmechanismen sind essenziell, um Vertrauen in KI-gestützte Prozesse aufzubauen und kontinuierliche Verbesserungen sicherzustellen.

7.4 Integration der KI-Thematik in den internen Auditplan

KI wird fester Bestandteil der internen Audits am IFP Basel und unterliegt einem standardisierten Prüfprozess.

- a) KI-Nutzung wird in den Management Review aufgenommen und bewertet.
- b) Alle neuen KI-Projekte durchlaufen eine Risikobewertung & Ethikprüfung durch den KI-Rat.
- c) Der KI-Nutzungsstatus wird quartalsweise überprüft und bei Bedarf angepasst.
- d) Die Ergebnisse der Audits werden dokumentiert & öffentlich gemacht.