



Checkliste – ab wann
liegt beim Finetuning
eine Entwicklung i.S.d.
KI-Verordnung vor?



1. Wie stark sind die Eingriffe in das KI-Modell im Rahmen des Finetunings?

Wichtig: Je tiefgreifender die Änderungen an deinem Modell (z. B. neue Layer, Modul-Integration, komplette Umbauten), desto höher ist die Wahrscheinlichkeit, dass du eine **erneute Konformitätsprüfung** brauchst.

	BEISPIELE	ERLÄUTERUNG
☐ Geringfügige Eingriffe	• Minimal anpassen: Du passt nur wenige Parameter an, etwa bei einer bestehenden Sentiment-Analyse, um sie leicht an deine eigene Wortwahl oder an einen Branchenjargon anzupassen. • Kleine Adapter-Schichten: Du fügst ein kleines „Adapter-Modul“ hinzu, damit dein Spam-Filter spezifische Firmenbegriffe erkennt – ansonsten bleibt das Grundmodell (z. B. ein gängiges Spam-Model) unverändert.	• Bei solch kleinen Veränderungen bleibt das Hauptmodell in seiner Struktur unangetastet. • Das Risiko, dass sich das Verhalten des gesamten Systems stark ändert, ist niedrig. • In der Praxis reicht hier oft eine einfache Dokumentation, da kein komplett neues Risiko-Profil entsteht.
☐ Mäßige Änderungen	• Zusätzliche Module integrieren: Du fügst mehrere Schichten hinzu, um deinen Textklassifizierer fachspezifisch zu trainieren – etwa für juristische Dokumente mit speziellen Paragraphen. • Erweiterte Fachsprache: Statt nur generelle Sentiment-Analyse auf Deutsch oder Englisch passt du das Modell auf seltene Fachbegriffe (z. B. medizinische Terminologie) an. Das Grundgerüst bleibt, aber du gehst über ein einfaches „Parameter-Tuning“ hinaus.	• Das KI-Modell hat erkennbare Änderungen, die über ein simples „Feintuning“ hinausgehen. • Die Struktur des Modells bleibt jedoch noch klar erkennbar und relativ ähnlich zum Original. • Ob hier eine erneute Konformitätsprüfung nötig ist, hängt stark von deinem Anwendungszweck ab: In Bereichen wie Medizin oder Recht ist die Risikobewertung meist genauer.
☐ Gravierende Modifikationen	• Kompletter Umbau: Du passt nicht nur Parameter an, sondern änderst massiv die Modellarchitektur (z. B. zusätzliche Encoder/-Decoder-Strukturen, deutlich mehr Layer). • Spezialisiertes Hochrisiko-Szenario: Du trainierst das Modell umfangreich für medizinische Diagnosen, wo eine minimale Fehlklassifizierung schwerwiegende Folgen haben kann – oder du wechselst von einer allgemeinen Bilderkennung zu einer spezialisierten MRT-Auswertung mit hochkomplexen Daten.	• Hier greifst du sehr tief in das Modell ein, sodass es in vielerlei Hinsicht fast „neu“ ist. • Alle bisherigen Annahmen zum Risikoprofil können sich ändern, da du wesentliche Kernbestandteile austauschst oder erheblich erweiterst. • In diesem Fall ist eine Prüfung praktisch immer wichtig, weil es sich rechtlich meist um eine „neue Entwicklung“ handelt.

2. Ergebnis / Anmerkungen

(Dokumentiere hier, was genau du änderst und in welchem Umfang. Beschreibe kurz, ob es sich eher um kleine Feineinstellungen oder um große Eingriffe handelt.)

Beispielhafte Dokumentation:

„Wir haben unser vortrainiertes Spam-Modell um eine zusätzliche Schicht erweitert, um typische Firmenjargons zu erkennen. Wir haben dazu 5.000 interne E-Mails gelabelt und das Modell 10 Epochen lang nachtrainiert. Die Architektur (BERT-Basis) blieb unangetastet.“

3. Prüfungsentscheidung

1. Keine wesentliche Änderung

- (Keine Entwicklung im Sinne der KI-Verordnung)
- Wenn du z. B. nur minimale Parameter anpasst (z. B. Anpassung einer vorhandenen Sentiment-Analyse für leicht andere Wortwahl), ist die Änderung in der Regel so geringfügig, dass keine neue Konformitätsprüfung nötig ist.

2. Eventuell wesentliche Änderung

- (Entwicklung denkbar; genauere Risikobewertung hinsichtlich Anwendungszweck nötig.)
- Trifft zu, wenn du über reine Feinjustierung hinausgehst, z. B. mehrere Layer hinzufügst, aber das Modell noch weitgehend erkennbar bleibt.
- Beispiel: Du erweiterst ein generisches Sprachmodell, um juristische Fachtermini zu erkennen – hier lohnt eine sorgfältige Prüfung, ob das neue Einsatzgebiet (Recht/Verträge) höhere Compliance-Anforderungen hat.

3. Wesentliche Änderung

- (Entwicklung liegt vor.)
- Wenn du die Modellarchitektur im Kern austauschst, das System auf eine andere Technologie umstellst oder in einem Hochrisikobereich (z. B. Medizin) komplett neue Funktionen integrierst.
- Beispiel: Du übernimmst ein allgemeines Bildklassifizierungsmodell und baust es so stark um, dass es eine seltene Krankheit in MRT-Bildern erkennt. In diesem Fall solltest du fast immer eine neue Konformitätsprüfung durchführen.