

Stiftung Bürgermut

# Code of Conduct Künstliche Intelligenz

---

Der Code of Conduct Künstliche Intelligenz (KI) hat zum Ziel, für die Mitarbeitenden der Stiftung Klarheit und Sicherheit im Umgang mit KI-Anwendungen zu schaffen und die Qualität der Arbeit zu sichern. Er soll den verantwortungsvollen Einsatz von KI innerhalb der Organisation fördern und gleichzeitig nach außen die ethischen Standards der Stiftung im Umgang mit dieser Technologie setzen. Der Code of Conduct wurde im Rahmen eines kollaborativen Prozesses zusammen mit den Mitarbeitenden entwickelt.

Unter „Künstlicher Intelligenz“ verstehen wir in diesem Dokument vor allem den Einsatz von generativen KI-Systemen wie Sprachmodellen (z. B. ChatGPT) zur Textverarbeitung, Bildgeneratoren zur visuellen Gestaltung sowie Werkzeugen zur Datenanalyse (z. B. OpenAI Data Analyst). Nicht gemeint sind klassische Automatisierungslösungen, Entscheidungsbäume oder regelbasierte Systeme.

Verantwortliche Ansprechperson für Fragen rund um die KI-Nutzung ist der KI-Beauftragte Henrik Flor. Er führt die Tool-Liste, organisiert den internen Wissenstransfer und die Aktualisierung dieses Codes.

## 1. Grundprinzipien

### **Zivilgesellschaftliche Verantwortung**

Als gemeinwohlorientierte Organisation ist ein kompetenter, reflektierter und transparenter Umgang mit KI Teil unserer gesellschaftlichen Vorbildrolle. Wir wollen eine überzeugende Balance zwischen Risiken und Nutzen finden.

### **Nachhaltigkeit und Ressourcenverbrauch**

KI benötigt große Rechenleistungen und damit auch Ressourcen. Ihr Einsatz soll daher vor allem dann erfolgen, wenn Arbeitsprozesse substanziell erleichtert und an deren Stelle Ressourceneinsparungen erreicht werden. Der Einsatz von KI ist eine bewusste, individuelle Abwägung jedes Mitarbeitenden.

### **Vorurteile und Verzerrungen (Biases)**

Wir sind uns bewusst, dass niemand frei von Vorurteilen ist. KI-Modelle können bestehende Biases in Trainingsdaten nicht nur übernehmen, sondern auch verstärken. Ein bewusster, kritischer Umgang mit diesen Risiken ist notwendig.

### **Datenschutz und Transparenz**

Datenschutz, Privatsphäre und Nachvollziehbarkeit von Entscheidungen sind wichtige Werte der Stiftung und finden auch beim Einsatz von KI-Modellen Anwendung.

### **Veränderungen im Team**

Wir beobachten und reflektieren aktiv, ob und wie KI unsere Zusammenarbeit im Team verändert. Soziale Beziehungen bleiben wichtiger als Effizienz.

### **Wissensaufbau und Teamkultur**

Wir fördern den sicheren und reflektierten Umgang mit KI durch regelmäßigen Austausch im Team und leicht zugängliche Wissensformate. Wir ermuntern Mitarbeitende auszuprobieren, Erfahrungen zu teilen und das neue Wissen zu dokumentieren.

### **Zivilgesellschaftlicher Austausch**

Wir tauschen uns mit anderen zivilgesellschaftlichen Akteur:innen zum Thema KI aus, teilen Wissen und machen dieses in verschiedenen Formaten anderen Organisationen niedrigschwellig zugänglich. Bewährte zivilgesellschaftliche KI-Anwendungen fördern wir aktiv.

## **2. Anforderungen**

### **Tool-Liste**

Die Stiftung stellt eine Positiv-Liste an Anwendungen zusammen, die Mitarbeitende nutzen können. Vorausgegangen ist eine Prüfung der Verlässlichkeit und möglicher Risiken. Die Liste ist für alle Mitarbeitenden im Stiftungs-Wiki auffindbar.

Weitere KI-Tools und KI-Apps dürfen unregistriert genutzt werden, wenn sie ein Impressum und eine eigene Datenschutzerklärung haben sowie transparent informieren, mit welchen Sprachmodellen sie arbeiten und wozu Daten genutzt werden. Die Verantwortung übernimmt der:die Nutzer:in. Für eine Registrierung mit einer Dienstadresse ist eine Benachrichtigung an bzw. Erlaubnis des KI-Beauftragten erforderlich.

### **Datenschutz**

Es werden keine personenbezogenen Daten in KI-Systeme eingespeist. Wir löschen oder anonymisieren solche Daten in Tabellen, Texten, Bildern und sonstigen Daten vor dem Upload.

In KI-Tools teilen wir keine vertraulichen oder sicherheitskritischen Daten. Das betrifft zum Beispiel Geschäftsgeheimnisse, interne Strategie-Papiere und sensible Konversationen. Wir schulen Mitarbeitende, damit alle KI datenschutzkonform anwenden.

### **Open Source**

Open-Source-Lösungen werden bevorzugt, sofern sie den Anforderungen genügen. Auf diese Weise soll die technologische Abhängigkeit von großen Unternehmen verringert werden.

### **Fakten-Check**

Quellenprüfung und Faktenchecks müssen erfolgen; es darf nicht vorausgesetzt werden, dass der Output von KI-Anwendungen vollständig oder objektiv richtig ist. Eine Überprüfung durch Menschen ist verpflichtend.

## **3. Qualitätsstandards**

### **Definition of Done**

KI-generierte Texte müssen der Definition of Done der Stiftung entsprechen, die im Scrum-Framework hinterlegt ist (vollständiger und fehlerfreier Inhalt, Faktencheck und Review, Dokumentation aktualisiert, Akzeptanzkriterien erfüllt).

### **Kennzeichnungspflicht**

An geeigneten Stellen, wie Datenschutzzinfos, Impressum, Artikel, geben wir die generelle Art der KI-Nutzung an (z. B. Erstellung, Redaktion, ko-kreatives Arbeiten, Illustration).

## **4. Nutzung**

### **Wofür wir KI-Anwendungen nutzen**

#### **Textbearbeitung**

Redigieren von Texten; Zusammenfassen und Kürzen von Headlines, Claims und Informationen; Unterstützung bei Recherchen.

#### **Konzeptionelle Arbeit**

Ideenentwicklung und Strukturierung von Inhalten. Die Stiftung nutzt KI als Sparringspartner:in zur Reflexion – nicht als generische:n Autor:in.

#### **Sprach- und Verständlichkeitsdienste**

Übersetzungen, Vereinfachung von Sprache, Erstellung von Untertiteln.

### **Dokumentation und Analyse**

Bearbeitung von Protokollen, einfache Datenanalysen oder Auswertung von Textinhalten, beispielsweise im Bereich Monitoring.

### **Sparring**

Bewertung/Analyse von Texten und anderen Dokumenten mit dem Ziel, deren Qualität zu verbessern.

### **Format-Flipping**

Konvertieren von Daten.

### **Barrierefreiheit**

Übersetzung von Texten in einfache Sprache.

## **Wofür wir KI-Anwendungen nicht nutzen**

### **Finanzielle Entscheidungen**

KI darf nicht für finanzielle Entscheidungen oder Budgetfragen herangezogen werden.

### **Bewerbungsprozesse**

KI darf nicht zur Auswahl oder Beurteilung von Bewerber:innen eingesetzt werden.

### **Ethisch sensible Vorgänge**

In Prozessen, bei denen ethische Überlegungen eine zentrale Rolle spielen, ist KI ausgeschlossen.

### **Faktisch nicht überprüfbare Inhalte**

Anwendungen, bei denen Ergebnisse nicht auf Richtigkeit oder Quellen zurückgeführt werden können, sind auszuschließen.

### **Bildgenerierung**

Bildgenerierung durch KI vermeiden wir, um die Authentizität der Bildsprache zu wahren und den Energieverbrauch gering zu halten. Ausnahmen sind nach Rücksprache im Team und mit entsprechender Kennzeichnung möglich.

### **Deepfakes und Bildfälschungen**

Der Einsatz von KI zur Erstellung manipulierter visueller Inhalte lehnt die Stiftung ab.

### **Protokolle**

Das automatische Generieren von Protokollen, etwa aus Videokonferenzen, nutzt die Stiftung nicht. Wir verstehen sie als geschützte digitale Räume. Inhalte sollen nicht Dienstleister:innen zur Verarbeitung zur Verfügung gestellt werden.

### **Verwendung agentischer KI**

Unter agentischer KI verstehen wir KI-Systeme, die selbstständig Ziele verfolgen, Entscheidungen treffen und Aufgaben in mehreren Schritten ausführen können. Diese verwendet die Stiftung bis auf weiteres nicht.

## **5. Verstöße gegen die Leitlinien**

Verstöße gegen die hier formulierten Grundsätze werden individuell durch den KI-Beauftragten und ggf. die Geschäftsleitung bewertet.

## **6. Weiterentwicklung Code of Conduct**

Der Code of Conduct ist ein *living document*, das – gerade mit Blick auf die schnelle technologische Weiterentwicklung dieses Bereichs – einer regelmäßigen Überprüfung bzw. Erweiterung bedarf.

Stand: 23. April 2026