

APA – AUSTRIA PRESSE AGENTUR

LEITLINIE ZUM UMGANG MIT KÜNSTLICHER INTELLIGENZ

Stand: März 2022



VORWORT

Künstliche Intelligenz (KI) hat das Potenzial, unsere Welt zu verbessern, birgt aber – wie viele an sich neutral wirkende Technologien – auch Risiken. Aus diesem Grund hat die Europäische Kommission 2021 einen neuen Rechtsrahmen für KI vorgelegt, der zum Ziel hat, das Konzept einer „vertrauenswürdigen KI“, die dem Menschen dient und menschenzentriert arbeitet, in Europa tief zu verankern. Damit soll der Einsatzbereich von künstlicher Intelligenz reguliert werden, insbesondere dort, wo Grundrechte und Datenschutz betroffen sind.

Im Medienbereich steht die Anwendung hochriskanter KI-Technologien im Normalfall nicht auf der Tagesordnung. Zusätzlich ist innerhalb der Datenschutzgrundverordnung (DSGVO) bzw. des Datenschutzgesetzes ein Medienprivileg zum Schutz der umfassenden Aufrechterhaltung der journalistischen Arbeit berücksichtigt. Doch ist es ein Gebot der Stunde, dass sich Medienunternehmen gezielt mit den Herausforderungen des neuen Rechtsrahmens auseinandersetzen. Denn zum einen stehen Medienunternehmen permanent im Blickpunkt der Öffentlichkeit und haben eine besondere Vorbildwirkung in Bezug auf ethische Standards. Zum anderen sind KI-Anwendungen im Alltag von RedakteurInnen, KuratorInnen, EditorInnen und Newsroom-DeveloperInnen angekommen, weshalb sie Leitlinien für deren Einsatz und Entwicklung benötigen.

Erste Ansätze, wie die APA – Austria Presse Agentur mit dieser Herausforderung umgehen könnte, wurden im Rahmen einer Studie des APA-medialab für das Technologieministerium (BMK) „AI.AT.Media“ im Sommer 2021 adressiert und in weiterer Folge in dem von der aws (Austria Wirtschaftsservice Gesellschaft m.b.H.) geförderten Projekt „Visual.Trust.AI“ im Rahmen von internen und extern begleiteten Workshops von einem cross-funktionalen Team ausgearbeitet. Das Ziel: Als innovatives, marktorientiertes Unternehmen möchte die APA die Vorteile von KI-Systemen nutzen, um Kundinnen und Kunden sowie Mitarbeiterinnen und Mitarbeiter bei ihrer Arbeit zu unterstützen.

Die vorliegende KI-Leitlinie zum Umgang mit künstlicher Intelligenz unterstreicht die Wahrnehmung der APA als rechtlich, technisch, inhaltlich und ethisch kompetente und innovative Partnerin. Sie stellt die Grundlage für alle Anwendungsgebiete künstlicher Intelligenz (Definition siehe S. 5) in der APA-Gruppe dar. Für den Anwendungsfall Gesichtserkennung in Bilddatenbanken, resultierend aus dem Projekt „Visual.Trust.AI“, sind hier bereits die Kernanforderungen formuliert.

Um den genannten Anforderungen zu genügen, orientiert sich die vorliegende KI-Leitlinie an den Ethik-Leitlinien für eine vertrauenswürdige KI der von der Europäischen Kommission eingesetzten, hochrangigen Expertengruppe für künstliche Intelligenz aus 2018. Unter Berücksichtigung der ethischen Grundsätze:

- ▶ Achtung der menschlichen Autonomie,
- ▶ Schadensverhütung,
- ▶ Fairness und
- ▶ Erklärbarkeit

strebt die APA die Einhaltung folgender Anforderungen an vertrauenswürdige KI bei der Entwicklung, Einführung und Nutzung von KI-Systemen an¹:

1. Vorrang menschlichen Handelns und menschliche Aufsicht
2. Technische Robustheit und Sicherheit
3. Schutz der Privatsphäre und Datenqualitätsmanagement
4. Transparenz
5. Vielfalt, Nichtdiskriminierung und Fairness
6. Gesellschaftliches und ökologisches Wohlergehen
7. Rechenschaftspflicht

¹ Vgl. Europäische Kommission (Hg.): Ethik-Leitlinien für eine vertrauenswürdige KI. Hochrangige Expertengruppe für künstliche Intelligenz, 2018.

INHALT

A.	GELTUNGSBEREICH	4
B.	DEFINITION VON KI	5
C.	ETHISCHE GRUNDSÄTZE FÜR APA-KI-SYSTEME	6
1.	ACHTUNG DER MENSCHLICHEN AUTONOMIE	7
2.	SCHADENSVERHÜTUNG	8
3.	FAIRNESS	9
4.	ERKLÄRBARKEIT	9
5.	SPANNUNGEN ZWISCHEN DEN GRUNDSÄTZEN	10

A. GELTUNGSBEREICH

Die gegenständliche KI-Leitlinie soll einen Handlungsrahmen für die Unternehmen der APA-Gruppe, konkret APA – Austria Presse Agentur eG und ihre Tochterfirmen APA-IT Informations Technologie GmbH, APA-DeFacto Datenbank & Contentmanagement GmbH, APA-OTS Originaltext-Service GmbH und Gentic Software GmbH bieten. Ist in Folge von „APA“ oder „APA-Konzern“ die Rede, sind darunter sinngemäß alle Unternehmen der APA-Gruppe zu verstehen, sofern nicht explizit die Redaktion der APA angeführt ist. Unter Berücksichtigung bestehender Regelwerke wie des APA-Redaktionshandbuchs oder des Ehrenkodex für die österreichische Presse des Österreichischen Presserats enthält die gegenständliche Leitlinie Handlungsempfehlungen zum Thema künstliche Intelligenz für die im Dokument beschriebenen Bereiche. Der Fokus liegt auf dem Einsatz und der Entwicklung von KI im Zusammenhang mit Produkten und Dienstleistungen der APA. Die Leitlinie dient als Handlungsrahmen für rechtlich und ethisch korrektes Verhalten. Gleichzeitig hilft er den Unternehmen dabei, Risiken aufzudecken und intern zu kommunizieren.

Es handelt sich um einen Handlungsrahmen, der sich laufend weiterentwickeln soll. Speziell in Hinblick auf die fortschreitende Digitalisierung von APA-Diensten und damit einhergehenden, neu entstehenden Rahmenbedingungen wird er einer periodischen Evaluierung unterzogen. Die Grundsätze einer vertrauenswürdigen KI, die in diesem Text detaillierter erklärt werden, stellen dabei das ethische Fundament für sämtliche aktuelle und künftige KI-Anwendungen der gesamten APA-Gruppe dar. Die darauf aufbauenden Kernanforderungen gehen auf Spezifika des jeweiligen KI-Anwendungsfalls ein und können sich von Fall zu Fall im Rahmen der Grundwerte unterschiedlich gestalten. Etwaige über den Rahmen der KI-Leitlinie hinausgehende Zusatzinformationen werden in separaten Dokumenten und Detailaufbereitungen zur Verfügung gestellt. Spätere Aktualisierungen werden bei entsprechender Relevanz im Unternehmen und auch nach außen (z. B. Öffentlichkeit, Förderstellen und weitere Stakeholder) kommuniziert.

Im Entstehen fokussiert sich die Leitlinie auf APA-Anwendungsfälle wie z. B. KI-Gesichtserkennung in internen Bilddatenbanken. Die aktuell bei der APA im Einsatz befindlichen KI-Anwendungen sind in erster Linie interne Systeme, welche Suchfilter bzw. Meta-Daten und/oder Texte bereitstellen. Überall dort, wo ein KI-gestütztes Endprodukt Kundinnen und Kunden zur Verfügung gestellt wird, werden die relevanten Punkte in den Kernanforderungen je Anwendungsfall berücksichtigt.

B. DEFINITION VON KI

KI-DEFINITION ALLGEMEIN

Eine allgemeingültige und konsistent genutzte Definition künstlicher Intelligenz existiert nicht. Am ehesten kann KI als Überbegriff für Anwendungen

genutzt werden, bei denen Maschinen menschenähnliche Intelligenzleistungen erbringen. Für die vorliegende KI-Leitlinie wird die Definition der Europäischen Kommission aus dem Jahr 2018 herangezogen:

„KI bezeichnet Systeme mit einem ‚intelligenten‘ Verhalten, die ihre Umgebung analysieren und mit einem gewissen Grad an Autonomie handeln, um bestimmte Ziele zu erreichen.“

„KI-basierte Systeme können rein softwaregestützt in einer virtuellen Umgebung arbeiten (z. B. Sprachassistenten, Bildanalysesoftware, Suchmaschinen, Sprach- und Gesichtserkennungssysteme), aber auch in Hardware-Systeme eingebettet sein (z. B. moderne Roboter, Drohnen oder Anwendungen des ‚Internet der Dinge‘).“²

KI-DEFINITION APA

In der APA fassen wir verschiedene KI-Definitionen folgendermaßen zusammen: Eine KI ist ein System, das in einem eng begrenzten Aufgabengebiet autonome, nicht-triviale Entscheidungen treffen kann. Sie bildet also bis zu einem gewissen Grad menschliche kognitive Fähigkeiten und Prozesse nach.

Als Voraussetzung für solche Entscheidungen muss eine KI immer an die konkreten Anforderungen jeder Aufgabenstellung angepasst werden. Diese Anpassung erfolgt in vielen Fällen durch maschinelles Lernen (Machine Learning). Machine Learning passiert immer unter kontrollierten Bedingungen und mit Hilfe von Trainingsdaten, welche für die Aufgabenstellung als geeignet ausgewählt und aufbereitet wurden. Durch diesen Prozess entsteht ein Modell, das Wissen über die Aufgabenstellung beinhaltet. Die Qualität des Modells wird durch Anwendung auf entsprechend gewählte Testdaten geprüft. Neben dem weiter oben bereits erwähnten Analysieren der Umgebung und

autonomen Handeln ist ein weiteres wichtiges Merkmal von KI also die Fähigkeit, aus Daten zu lernen und dadurch ihr Verhalten anzupassen.

VERTRAUENSWÜRDIGE KI

Eine weitere Spezifizierung stellt jene der „vertrauenswürdigen KI“ dar, der sich die vorliegende Leitlinie verpflichtet fühlt. Vertrauenswürdige KI zeichnet sich durch drei Komponenten aus, die während des gesamten Lebenszyklus des Systems erfüllt sein sollten:

1. Die KI sollte **rechtmäßig** sein und somit geltendes Recht und alle gesetzlichen Bestimmungen einhalten. Dazu gehören folgende Grundrechte: Achtung der Menschenwürde, Freiheit des Einzelnen, Achtung von Demokratie, Gerechtigkeit und Rechtsstaatlichkeit, Gleichheit, Nichtdiskriminierung und Solidarität, Bürgerrechte.
2. Sie sollte **ethisch** sein und somit die Einhaltung ethischer Grundsätze – Achtung der menschlichen Autonomie, Schadensverhütung, Fairness, Erklärbarkeit und Werte – garantieren; und
3. sie sollte **robust** sein, und zwar sowohl in technischer als auch in sozialer Hinsicht, da KI-Systeme möglicherweise unbeabsichtigten Schaden verursachen, selbst wenn ihnen gute Absichten zugrunde liegen.

² Vgl. Bundesministerium für Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie (Hg): AI.AT.MEDIA, AI and the Austrian Media Sector: Mapping the Landscape, Setting a Course, 2021. https://www.bmk.gv.at/themen/innovation/publikationen/ikt/ai/ai_at_media.html

RISIKOEINSCHÄTZUNG

Von der KI-Leitlinie erfasst sind prinzipiell alle von der APA im Bereich Produkte und Services entwickelten Systeme, heute und in Zukunft. Der Verordnungsentwurf der Europäischen Kommission unterteilt KI-Systeme in unterschiedliche Klassen ³:

- ▶ **Hochrisiko-KI-Systeme:** Darunter versteht die EU KI-Systeme, bei denen es aus Sicht der Kommission besonders wahrscheinlich ist, dass ein Eingriff in Sicherheit, Gesundheit oder Grundrechte stattfindet. Für diese Systeme ist eine Zertifizierung nach einheitlichen EU-Standards notwendig. Ob ein KI-System ein Hochrisiko-KI-System ist, hängt nicht von dessen Ausgestaltung, sondern von dessen Zweckbestimmung ab.
- ▶ **Geringes Risiko:** An KI-Systeme, die nicht als „hochrisikobehaftet“ gelten, aber der Interaktion mit natürlichen Personen dienen (z. B. Chatbots), werden besondere Transparenzanforderungen gestellt. Den Nutzerinnen und Nutzern soll klar sein, dass sie es mit einem KI-System zu tun haben. Wer etwa ein Emotionserkennungssystem einsetzt oder ein System zur Assoziierung (gesellschaftlicher) Kategorien anhand biometrischer Daten, muss die davon

betroffenen natürlichen Personen über den Betrieb des Systems informieren. KI-Systeme, die sogenannte Deepfakes erzeugen, müssen offenlegen, dass diese Inhalte künstlich erzeugt bzw. manipuliert wurden. Ein Deepfake ist ein Bild-, Ton- oder Videoinhalt, der einer wirklichen Person, einem Gegenstand oder Ereignis merklich ähnelt und fälschlicherweise als echt oder wahrhaftig eingestuft werden könnte.

- ▶ **Minimales Risiko:** Systeme mit minimalem Risiko können weitgehend ohne Beschränkungen (im Rahmen des sonstigen Rechts) verwendet werden. Beispiele wären etwa Videospiele oder Spamfilter.

Um den Anforderungen an KI in den Bereichen „Hochrisiko“ und „geringes Risiko“ auch bei künftigen Entwicklungen gerecht werden zu können, verpflichtet sich die APA zu einer internen Prüfung jeder neuen KI-Anwendung in Hinblick auf deren Einstufung. Die Prüfung findet entsprechend der Handlungsempfehlungen der KI-Leitlinie durch die verantwortliche Person statt, welche die Kernanforderungen der gegenständliche KI erarbeitet. Im Falle einer Einstufung über „Geringes Risiko“ hinaus ist ein rechtliches Gutachten einzuholen, um das potenzielle Vorhandensein einer Hochrisiko-KI eindeutig klären zu können.

C. ETHISCHE GRUNDSÄTZE FÜR APA-KI-SYSTEME

Neben Qualität, Sicherheit und Ausgewogenheit stellt Vertrauenswürdigkeit einen zentralen Wert der APA-Redaktion, aber auch der gesamten APA-Gruppe dar. Sie hat die Kernaufgabe, „Trusted Information und Trusted Content“ nach Qualitätskriterien (Check, Double-Check, Re-Check) für Medien zu produzieren. Diese und weitere strenge Qualitätskriterien für journalistisches Arbeiten (siehe Seite 8) gelten auch für in die redaktionelle Arbeit eingebundene KI-Anwendungen. Um die Ver-

trauenswürdigkeit des KI-Systems sicherzustellen, setzt die APA Content-Beschränkungen (nur ausgesuchte, geprüfte Quellen) oder Nutzergruppen-Beschränkungen (nur ausgewählte Nutzerinnen und Nutzer haben Zugriff auf die Daten) ein.

Diese Entscheidungen (welche Quellen für welche Nutzergruppen) werden transparent dargestellt. Im Einzelfall können auch beide Methodiken Grundlage einer KI-Anwendung sein.

³ Vgl. Kieffer, Dr. Michael: CE-Zeichen für KI-Systeme – Ausweitung des Produktsicherheitsrechts auf künstliche Intelligenz, 2021. Siehe <https://iot.taylorwessing.com/ce-zeichen-fuer-ki-systeme-ausweitung-des-produktsicherheitsrechts-auf-kuenstliche-intelligenz>

Die APA verpflichtet sich, bei der Entwicklung, Einführung und Nutzung von KI-Systemen folgende ethische Grundsätze einzuhalten:

1. Achtung der menschlichen Autonomie
2. Schadensverhütung
3. Fairness
4. Erklärbarkeit

1. ACHTUNG DER MENSCHLICHEN AUTONOMIE

1.1. ACHTUNG DER GRUNDRECHTE

Das Fundament für eine vertrauenswürdige KI stellen die Grundrechte dar. Die APA sieht es als ihre rechtliche und ethische Verantwortung an, dass ihre KI-Systeme im Bereich Produkte und Services mit den folgenden Grundrechten in Einklang stehen:

- ▶ Achtung der Menschenwürde,
- ▶ Freiheit des Einzelnen,
- ▶ Achtung von Demokratie, Gerechtigkeit und Rechtsstaatlichkeit,
- ▶ Gleichheit, Nichtdiskriminierung und Solidarität,
- ▶ Bürgerrechte

Datenschutz und Privatsphäre nehmen dabei einen hohen Stellenwert ein⁴. Potenziell negative Auswirkungen von KI auf Grundrechte werden einer gewissenhaften Vorab-Prüfung unterzogen und in Hinblick auf die maximal mögliche Schadensverhütung behandelt.

1.2. JOURNALISTISCHE QUALITÄTSSTANDARDS

Die im APA-Statut verankerten Qualitätskriterien Unabhängigkeit, Objektivität, Ausgewogenheit, Vermeidung von Einseitigkeit und Parteinahme, Zuverlässigkeit, Richtigkeit und Schnelligkeit stellen die Grundlage sämtlicher journalistischer Tätigkeit der Redaktion der APA dar.⁵

Zudem hat sich die Redaktion der APA dem Ehrenkodex für die österreichische Presse des Österreichischen Presserats verpflichtet. Durch die Einhaltung der im Grundsatzkatalog festgelegten Inhalte Freiheit, Genauigkeit, Unterscheidbarkeit, Vermeidung von Einflussnahmen, Persönlichkeitsschutz, Intimsphäre, Schutz vor Pauschalverunglimpfungen und Diskriminierung, Vermeidung unlauterer Methoden bei der

Materialbeschaffung, redaktionelle Spezialbereiche, öffentliches Interesse, Interesse von MedienmitarbeiterInnen sowie Suizidberichterstattung bekennt sich die Redaktion der APA zu einem der Wahrheitsfindung und Korrektheit verpflichteten Gebrauch der Pressefreiheit.⁶

Die Redaktion und ihr Wertegerüst wiederum bilden als Kern der Nachrichtenagentur die Basis aller mit der Marke APA in Verbindung stehenden Entitäten. Sämtliche Dienste und Produkte der APA-Gruppe, so auch KI-basierte Lösungen, orientieren sich bei entsprechender Anwendbarkeit an diesen hohen redaktionellen Qualitätsstandards. Innovationen im Bereich der Quellenverarbeitung, Informationsaufbereitung, Contentverarbeitung und -distribution werden durch Re-Evaluierung dieser Standards begleitet, Qualitätssicherungsprozesse gegebenenfalls für neue Anwendungen in diesen Bereichen adaptiert.

1.3. INTERNATIONALE BRANCHENSTANDARDS

Im Rahmen der Zusammenarbeit mit anderen Nachrichtenagenturen auf bilateraler, europäischer (Gruppe 39, EANA – European Alliance of News Agencies) und globaler Ebene (MINDS International) findet ein kontinuierlicher Austausch zu Themen aller Art – so auch Innovationsthemen wie KI – statt. Unabhängig ihrer geografischen Verortung und (nicht-)kommerziellen Orientierung stellen Redaktionen mit ihren hohen journalistischen Branchenstandards den Kern jeder Nachrichtenagentur dar. Auch auf internationaler Ebene bildet somit die Einhaltung der Branchenstandards die Grundlage sämtlicher Kooperationen und gemeinschaftlicher Entwicklungen von Produkten und Dienstleistungen.

⁴ Vgl. APA – Austria Presse Agentur: Datenschutzerklärung APA-Gruppe, Stand Mai 2021, siehe <https://apa.at/about/datenschutzerklaerung>

⁵ vgl. APA – Austria Presse Agentur: APA-Redaktionshandbuch, 2005, siehe <https://apa.at/about/die-redaktion>

⁶ Österreichischer Presserat (2019): Grundsätze für die publizistische Arbeit (Ehrenkodex für die österreichische Presse), siehe https://www.presserat.at/show_content.php?hid=1

2. SCHADENSVERHÜTUNG

Von der APA erstellte KI-Systeme sollen weder Schaden verursachen noch verschärfen oder sich auf andere Art und Weise negativ auf Menschen auswirken. Der Schutz der Menschenwürde sowie der geistigen und körperlichen Unversehrtheit steht im Vordergrund. In erster Linie sollen KI-Systeme dazu dienen, Arbeitsschritte zu vereinfachen, vorhandene Daten zu ergänzen, diese besser aufzubereiten bzw. besser auffindbar zu machen.

Die APA setzt all ihre Handlungen gemäß den gültigen gesetzlichen Rahmenbedingungen. Vor dem Einsatz einer KI-Anwendung werden daher die nötigen rechtlichen Prüfungen, z. B. hinsichtlich Persönlichkeitsrechte, vorgenommen, um materielle wie immaterielle Schäden einer Person zu verhindern.

Zusätzlich werden Produkte und Services in der Regel entlang der bereits erwähnten hohen journalistischen Standards entwickelt. Der Ausprägungsgrad variiert je Kundengruppe und Anwendungsfall, wird jedoch in den Kernanforderungen zur jeweiligen KI-Lösung getrennt erörtert.

Um diesen Anforderungen gerecht werden zu können, werden die Anwendungsfälle für den Einsatz von KI so eng und spezifisch wie möglich definiert und die technischen Möglichkeiten nicht zwingend zur Gänze ausgeschöpft. Bewusste Beschränkungen zum Zwecke der Schadensverhütung können – je nach KI-Anwendung und Umsetzbarkeit – Content- und/oder Nutzungs-seitig erfolgen.

In allen Stufen der konzeptionellen Entwicklung von KI-Systemen der APA-Gruppe werden systemische Sicherungsmaßnahmen nach technischer Möglichkeit getroffen. APA-Tech als technischer Dienstleister des gesamten APA-Konzerns wurde von TÜV AUSTRIA nach ISO/IEC 20000-1:2018 sowie nach ISO/IEC 27001:2013 zertifiziert, wobei ISO/IEC 20000-1:2018 die Einhaltung der Standards für das IT-Service-Management belegt und ISO/IEC 27001:2013 ein Qualitätssiegel für Informationssicherheit darstellt. Letzteres bescheinigt die Einhaltung der hohen Qualitätslevels für

IT-Security Management und stellt sicher, dass für den Basisbetrieb, der hinter den technischen Serviceleistungen von APA-Tech steht, alle relevanten Maßnahmen bezüglich Datensicherheit, Rechtssicherheit, Integrität und Verfügbarkeit umgesetzt werden.

Die APA-Tech Cyber Security Policy fasst wesentliche, vereinbarte, anzuwendende Grundsätze und Richtlinien zum sicheren Umgang mit IT-Systemen und Daten für die APA-Gruppe zusammen, die detailliert in internen Richtlinien, Prozessbeschreibungen, Betriebs- oder Dienstvereinbarungen geregelt sind.

Zusätzlich kommt auf menschlicher Seite je nach Sensibilität das journalistische Vier-Augen-Prinzip als wesentliches Kriterium des Qualitätsmanagements und der Überwachung bzw. der Schadensverhütung zum Einsatz. Ein Beispiel für eine solche kritische Anwendung ist die KI-Gesichtserkennung in Bilddatenbanken, hier werden die hinterlegten Trainingsdaten im 4-Augen-Prinzip geprüft, da sie die Grundlage der Funktionalität bilden. Dies gilt vor allem für die Grundlagen der KI-Anwendung wie zum Beispiel Trainingsdaten.

2.1. NACHHALTIGKEIT

Beim Training und Einsatz von KI-Systemen achten wir darauf, für jede Aufgabenstellung eine möglichst effiziente Systemarchitektur zu entwerfen. Wir entwickeln und nutzen also stets auf konkrete Aufgaben spezialisierte Ansätze.

Durch die Spezialisierung sind die von uns genutzten Modelle um viele Größenordnungen kleiner als die häufig in den Medien für ihren Ressourcenbedarf kritisierten allgemeinen KI-Modelle. Außerdem streben wir an, jedes Modell auf der dafür am besten geeigneten Hardware auszuführen.

Durch diese Grundsätze minimieren wir den Ressourceneinsatz und daraus folgende ökologische Auswirkungen.

3. FAIRNESS

Die APA verpflichtet sich, bei der Entwicklung von KI-Systemen den Grundsatz der Fairness stets zu berücksichtigen. Das Ziel dabei ist eine gleiche und gerechte Verteilung von Vorteilen und Kosten sowie die Gewährleistung des Schutzes von Personen und Gruppen vor unfairer Verzerrung, Diskriminierung und Stigmatisierung. Die APA steuert möglichen Verzerrungen entgegen, indem Bias in den Daten für das Training von KI so weit wie möglich vermieden wird, etwa durch zusätzliche Recherchen.

Sollte ein gewisser Bias in einer KI-Anwendung absolut unvermeidbar sein, so wird dieser seitens APA transparent kommuniziert. Auf die Gefahr von Diskriminierung oder Stigmatisierung wird bereits bei der Produktentwicklung Bedacht genommen. Darüber hinaus wird die APA dort, wo es notwendig ist, an das jeweilige Produkt oder Service angepasste Opt-In- oder Opt-Out-Möglichkeiten anbieten (z. B. Einholen von Einverständniserklärungen oder das Löschen aus einem Index).

4. ERKLÄRBARKEIT

„Ohne menschliche gibt es keine künstliche Intelligenz. Nur wenn wir verstehen, was die Maschinen machen, können wir sie auch so einsetzen, dass es Sinn macht.“⁷

Die APA verpflichtet sich, Prozesse im Zusammenhang mit KI-Anwendungen im größtmöglichen Ausmaß transparent zu gestalten und Rückverfolgbarkeit sowie Erklärbarkeit zu gewährleisten, soweit dies technisch realisierbar ist. Damit sollen neue Standards wie IEEE 7000 erreicht werden. Dies umfasst etwa die sorgfältige Dokumentation der in die KI-Anwendung involvierten Prozesse, Mechanismen und Entscheidungsabläufe, der zugrunde liegenden Trainings- und Testdaten, der Evaluierung des Modells, der Ergebnisse interner Testläufe sowie der finalen Validierung. Ebenso umfasst sind neben der Entwicklungs- und der Einführungsphase auch die Ergebnisse der laufenden Evaluierung von KI-Anwendungen im Echtbetrieb wie die stichprobenartige Durchsicht von Ergebnissen oder die quantitative Analyse von Suchabfragen.

In manchen Fällen ist die APA zum Erreichen der bestmöglichen Resultate auf die Nutzung externer Dienstleister angewiesen. Die Wahl eines Dienstleisters basiert neben technischen Eigenschaften stark auf seiner Transparenz in Bezug auf die im vorigen Absatz genannten Aspekte. Auch bei Nutzung externer Dienstleister verwendet die APA so weit wie möglich Trainings- und Testdaten, die von ihr selbst zusammengestellt und aufbereitet wurden.

Die Fähigkeiten und Limitationen von KI-Anwendungen werden seitens APA so transparent wie möglich kommuniziert, um das dauerhafte Vertrauen der NutzerInnen in KI-Anwendungen des Unternehmens gewährleisten zu können. Die Kommunikation bedient sich je nach KI-Anwendung sämtlicher relevanter Kanäle wie Webseiten und Plattform-Oberflächen und gestaltet sich – im Einklang mit dem lebenden Dokument der KI-Leitlinie – adaptiv.

⁷ APA-DeFacto Geschäftsführer Klemens Ganner am Österreichischen Kommunikationstag 2021 des PRVA.

5. SPANNUNGEN ZWISCHEN DEN GRUNDSÄTZEN

Die ethischen Grundsätze wie Achtung der menschlichen Autonomie, Schadensverhütung, Fairness und Erklärbarkeit können miteinander in Konflikt geraten. Am Beispiel KI-gestützter Gesichtserkennung zeigt sich, dass diese einerseits eingesetzt werden kann, um Kriminalität zu reduzieren (Schadensverhütung), andererseits aber auch das Potenzial hat, die Privatsphäre und Freiheitsrechte von Individuen (Achtung der menschlichen Autonomie) einzuschränken. In solchen Fällen wird mittels faktenbasierter Reflexion ein vernünftiger Kompromiss angestrebt. Die Prämisse, dass die Vorteile

von KI-Systemen gegenüber den vorhersehbaren, individuellen Risiken erheblich überwiegen sollen, wird stets berücksichtigt. Sollte ein Kompromiss nicht möglich sein, ist absoluten Prinzipien wie der Würde des Menschen Vorrang gegenüber konkurrierenden Grundsätzen einzuräumen. Vorgänge dieser Art sind kontinuierlich zu ermitteln, bewerten, dokumentieren und zu kommunizieren.

MEDIENINHABER, HERAUSGEBER UND HERSTELLER:

APA – Austria Presse Agentur
Laimgrubengasse 10, 1060 Wien, Österreich
+43 1 36060
apa@apa.at
www.apa.at

IDEE UND KONZEPT:

Kathrin Karner, Luzia Strohmayer-Nacif,
Verena Krawarik, Bernhard Sonntag

REDAKTION:

Kathrin Karner, Luzia Strohmayer-Nacif, Verena Krawarik,
Bernhard Sonntag, Marian Rainer-Harbach, Michael Paulhart,
Irene Korinek, Andreas Mauczka, Katharina Schell

GESTALTUNG UND SATZ:

vektorama. grafik.design.strategie OG

DATUM:

März 2022

ERSCHEINUNGSORT:

Wien