



DEZA-Arbeitshilfe

zur künstlichen Intelligenz (KI)

(Dokumenttyp C)

Genehmigt an der Klausur der DEZA-Direktion vom 27.06.2025



Lizenzfreies Foto, heruntergeladen von: <https://www.vecteezy.com/>¹.

Inhalt

1. Warum ist Künstliche Intelligenz (KI) wichtig für die DEZA.....	2
2. Definition und gesetzlicher & politischer Rahmen.....	3
3. Künstliche Intelligenz in der DEZA: SWOT-Analyse	4
4. 10 DEZA-Leitsätze für einen verantwortungsvollen KI-Einsatz	6
5. Die DEZA als Geldgeberin für KI (Finanzierung von operationellen Projekten)	6
6. Die DEZA als Kooperationspartnerin im KI-Bereich (Einflussnahme auf die Politik)	8
7. Die DEZA als Beraterin über KI (thematische Beratung)	10
8. Die DEZA als Wissensvermittlerin durch KI (Wissensmanagement).....	10
9. KI in DEZA-internen Prozessen.....	11
 <i>Glossar</i>	 12

Dokument erstellt von der Sektion Frieden, Gouvernanz und Gleichstellung der DEZA im Rahmen des Mandats «*PGE: thematische Unterstützung der Gruppe/Sektion PGE*», das dem Konsortium **P4PGE** übertragen wurde. Für P4PGE: Caroline Khene (C.Khene@ids.ac.uk), Tony Roberts (T.Roberts@ids.ac.uk) und Kevin Hernandez (k.hernandez2@ids.ac.uk), alle vom *Institute of Development Studies UK*, Library Road, Brighton BN1 9RE, Vereinigtes Königreich.

1. Warum ist Künstliche Intelligenz (KI) wichtig für die DEZA

Künstliche Intelligenz (KI) wird sich weltweit mindestens genauso stark auf das Privat- und das Berufsleben der aktuellen und der kommenden Generationen auswirken wie die Einführung des Internets für die breite Öffentlichkeit in den 1990er-Jahren. Auch die Direktion für Entwicklung und Zusammenarbeit (DEZA) wird nicht darum herumkommen, die Stärken und Möglichkeiten von KI-Technologien zu nutzen. Gleichzeitig muss sie einen geeigneten Umgang mit den Schwächen und Risiken von KI finden (siehe Kapitel 3: SWOT-Analyse).

Warum sollte KI für die DEZA wichtig sein



Bessere Entscheidungsfindung durch Dateneinblicke

Die KI kann große Datensätze analysieren, um Trends, Risiken und Chancen zu erkennen und so evidenzbasierte Politikplanung und Wirkungsprognosen ermöglichen.



Höhere Effizienz und Kosteneffektivität

Die KI kann Routinetätigkeiten wie Berichterstattung und Übersetzung automatisieren, wodurch Mitarbeiter für strategische Aufgaben frei werden und die Verwaltungskosten senken.



Gezielte Hilfeleistung

Maschinelles Lernen kann Verwundbarkeiten vorhersagen, Armut kartieren und Klimafolgen in Echtzeit überwachen – für eine präzisere und zeitgerechte Hilfeleistung.



Stärkung lokaler Resilienz und Innovation

KI-Tools können lokale Akteure z.B. in klimaintelligenter Landwirtschaft, Telemedizin oder KI-gestützten Marktzugangsplattformen bevollmächtigen.



Ethische Führung und globale Standards

Die DEZA kann sicherstellen, dass KI ethisch in der internationalen Zusammenarbeit eingesetzt wird – durch Förderung von Transparenz, Fairness und Rechenschaftspflicht weltweit.

Alle Länder und Bereiche, alle Abteilungen und Sektionen der DEZA werden sich nach und nach mit den KI-Technologien auseinandersetzen müssen. Deshalb **gilt die vorliegende Arbeitshilfe für die gesamte DEZA.**

Die Schweiz verfügt über einen rechtlichen und politischen Rahmen mit einer anerkannten Definition von KI (Kapitel 2). KI beinhaltet sowohl positive als auch negative Aspekte, die für die DEZA relevant sind (Kapitel 3). Daraus lassen sich Grundsätze für eine verantwortungsvolle KI-Nutzung ableiten (Kapitel 4). Zudem bestehen Richtlinien für alle Schlüsselrollen der DEZA: als Geldgeberin für Projekte und Programme (Kapitel 5), als Kooperationspartnerin, die Einfluss auf die Politik nimmt (Kapitel 6), als Beraterin der verschiedenen Organisationseinheiten und der Vertretungen (Kapitel 7), als Wissensvermittlerin (Kapitel 8), einschliesslich des Aufbaus von KI-Kapazitäten bei den DEZA-Mitarbeitenden. Auch bei den DEZA-internen Abläufen spielt KI eine Rolle (Kapitel 9).

Diese Arbeitshilfe wird durch separate, dynamische Dokumente ergänzt, die unabhängig von Direktionsentscheiden aktualisiert werden. So bleibt die Arbeitshilfe mit präzisen, praktischen Anleitungen, Beispielen und Schulungsmaterialien auf dem neusten Stand.

2. Definition und gesetzlicher und politischer Rahmen

Am 27. März 2025 hat die Schweiz das rechtsverbindliche **Rahmenübereinkommen des Europarats über künstliche Intelligenz und Menschenrechte, Demokratie und Rechtsstaatlichkeit (05.09.2024)** unterzeichnet². Das Übereinkommen soll dafür sorgen, «dass alle Tätigkeiten im Lebenszyklus von Systemen der künstlichen Intelligenz uneingeschränkt mit den Menschenrechten, der Demokratie und der Rechtsstaatlichkeit vereinbar sind. Jede Vertragspartei trifft geeignete Rechts-, Verwaltungs- oder sonstige Massnahmen oder erhält diese aufrecht, um den in diesem Übereinkommen festgelegten Bestimmungen Wirkung zu verleihen.» (Art. 1).

Definition gemäss Rahmenübereinkommen (Art. 2):

«System der künstlichen Intelligenz» bezeichnet «ein maschinengestütztes System, das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können.»

KI ist damit weder künstlich noch intelligent, sondern eine Reflexion aus kollektiven Daten, die durch Algorithmen projiziert werden.

Für die Schweiz gilt zudem das **Bundesgesetz über den Datenschutz**³, das mit der *Datenschutz-Grundverordnung der EU* in Einklang steht⁴. Die wichtigsten KI-bezogenen Bestimmungen sind in Artikel 21 enthalten. Demnach müssen betroffene Personen über Entscheidungen informiert werden, die ausschliesslich auf einer automatisierten Bearbeitung beruhen und die für sie mit einer Rechtsfolge verbunden sind oder sie erheblich beeinträchtigen.

Als Gründungsmitglied der 1961 ins Leben gerufenen Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) hat die Schweiz auch den **OECD-Grundsätzen für künstliche Intelligenz**⁵ vom Mai 2019 zugestimmt, die im Mai 2024 aktualisiert wurden und Teil der umfassenderen KI-Gouvernanz sind. Die Schweiz ist Mitglied des DAC Gouvernanz Network (GovNet) und hat 2025 dessen Vorsitz inne. In diesem Rahmen war die Schweiz 2025 massgeblich an der Schaffung eines neuen Arbeitsbereichs zur digitalen Transformation beteiligt, der auch KI einschliesst. Im Jahr 2021 unterstützte die Schweiz zudem die **UNESCO-Empfehlung zur Ethik der künstlichen Intelligenz**⁶. In der **Strategie der internationalen Zusammenarbeit der Schweiz**⁷ werden neue Technologien, darunter KI, als **zentrale Modalitäten/Ansätze** der internationalen Zusammenarbeit bezeichnet. Ausserdem bestehen innerhalb der Bundesverwaltung **Vorgaben für die Verwendung von KI**⁸.

3. Künstliche Intelligenz in der DEZA: SWOT-Analyse

Als Teil des Eidgenössischen Departements für auswärtige Angelegenheiten (EDA) und damit der Bundesverwaltung ist die DEZA an den genannten rechtlichen und politischen Rahmen gebunden. Die DEZA unterstützt Entwicklungsländer unter Berücksichtigung des Kontexts des jeweiligen Partnerlandes bei der Verbesserung der Lebensbedingungen für die Bevölkerung⁹. Die **Stärken, Schwächen, Chancen und Risiken (SWOT) der Nutzung von KI-Technologien in der internationalen Zusammenarbeit** lassen sich für die DEZA wie folgt zusammenfassen:

<u>STÄRKEN</u>	<u>SCHWÄCHEN</u>
<ol style="list-style-type: none">1. Prädiktive Fähigkeiten und Prognosen zur Unterstützung einer raschen Entscheidungsfindung und von Strategien als Reaktion auf ungewisse oder neue Kontexte, gestützt auf Ausgangsdaten, die Verhaltensmuster und kontextuelle Veränderungen erkennen.2. Bestandsaufnahme und Identifizierung für Situationsanalysen, um Einsätze vorzubereiten und möglichst wirksame oder effiziente humanitäre Massnahmen zu planen.3. Lehren aus verpassten Mustern in historischen und aktuellen Daten ziehen, um die aktuelle Praxis zu gestalten und zu beraten.4. Übersetzung und Kommunikation: KI-Unterstützung erleichtert die Arbeit in einem multikulturellen Kontext erheblich.5. Aufgaben und Austausch automatisieren, wobei für Routineaufgaben KI eingesetzt wird, damit sich die Mitarbeitenden auf wichtigere Aufgaben konzentrieren können und durch KI-Interaktionen ein kontinuierlicher Austausch mit der Bevölkerung bzw. den Gemeinschaften sichergestellt ist.	<ol style="list-style-type: none">1. Risiko, dass die digitale Kluft zum Globalen Süden durch die Hilfsmassnahmen zunimmt, weil der Zugang zu grundlegender digitaler Infrastruktur, zu Ressourcen, Gouvernanz, integrativen Ansätzen, KI-Fähigkeiten und Bildung (digitale Kompetenzen) fehlen, die für die wirksame Nutzung von KI-Technologien erforderlich sind, wodurch bestehende sozio-ökonomische Ungleichheiten und Entwicklungsunterschiede noch verschärft werden. Daten und Berechnungen/Computing sind wichtige KI-Elemente, die eine umfassende, kontextbezogene KI-Praxis einschränken.2. Herausforderung, generative KI-Systeme (siehe Definition im Glossar) mit dem vielfältigen und komplexen Spektrum menschlicher Werte, Überzeugungen und ethischer Grundsätze in verschiedenen Kulturen, Sprachen und normativen Rahmen in Einklang zu bringen und soziale Neutralität zu wahren, indem diskriminierende oder kulturell unsensible Ergebnisse vermieden werden, insbesondere in Bereichen wie der politischen Entscheidungsfindung.3. Schwächen in Technik und Cybersicherheit: Eingaben können gespeichert, von den Systemen wiederverwendet und in späteren Abfragen reproduziert werden. Unberechtigte Benutzerprofile können erstellt und Eingabedaten missbraucht werden. Gelegentlich «halluzinieren» die Systeme und liefern einen völlig falschen Output. Je nach Sprachmodell kann die Datenbasis veraltet und der Output falsch sein, was schädliche Folgen haben kann, insbesondere in der Medizin oder in der Beratung.4. Fehlende Rechenschaftspflicht: Das Verhalten fortgeschrittener KI-Systeme ist kaum nachvollziehbar, was es schwierig macht, potenzielle Risiken zu erkennen und die Verantwortung für das Risikomanagement zuzuordnen.5. Datendiebstahl, Datenkolonialismus und Arbeitsausbeutung: Fehlende Einflussnahme und die Machtdynamik zwischen den Hilfsakteuren und der betroffenen Bevölkerung bei der Entwicklung und beim selektiven Einsatz von KI-Systemen nähren den digitalen Kolonialismus, indem Daten von gefährdeten Bevölkerungsgruppen ohne Möglichkeit zur Einflussnahme oder Zustimmung gesammelt und aus der Ferne verarbeitet werden. Diesen Bevölkerungsgruppen wird die Autonomie über ihre digitalen Darstellungen und Informationen verwehrt. Wenn gesammelte Daten im Globalen Süden verarbeitet werden, dann oft in Ländern mit niedrigen Löhnen und lückenhaften Arbeits- und Datenschutzbestimmungen.6. Wenig transparente Algorithmen angesichts der undurchsichtigen Beschaffenheit der KI (Blackbox), die es schwierig macht, gegebenenfalls willkürliche Entscheidungsergebnisse nachzuvollziehen. Noch schwieriger wird es bei Systemen des maschinellen Lernens (siehe Definition im Glossar), bei denen ein kontinuierlicher Lernprozess stattfindet. Programme in sensiblen Bereichen wie Frieden, Menschenrechte, Gouvernanz und soziale Inklusion können durch falsche und diskriminierende automatisierte Entscheidungen beeinflusst werden.7. Mangelnde Kompetenzen und Fähigkeiten der Mitarbeitenden der DEZA/EDA oder von Partnern: Da KI-Technologien für die breite Öffentlichkeit relativ neu sind (2022) und der Bund und das EDA erst seit Kurzem über erste Regelungen verfügen, stehen den Verwaltungsmitarbeitenden aktuell nur begrenzte Kenntnisse und Schulungen zur Verfügung.8. Wichtiger klimatischer/ökologischer Fussabdruck: KI-Technologien benötigen hohe Strommengen, die häufig aus nicht erneuerbaren Energiequellen stammen, was erhebliche CO₂-Emissionen verursacht.

CHANCEN

1. Die Schweiz verfügt über ein stabiles und günstiges Umfeld für Innovationen im KI-Bereich, das sich auf Hochschulen, Forschungseinrichtungen und innovative Privatunternehmen (weltweit führend bei den KI-Start-ups pro Kopf) stützt, sowie über **einen der wenigen Supercomputer¹⁰ der Welt, mit privilegiertem Zugang für Initiativen des Globalen Südens¹¹.**
2. International / von der EU¹² / vom Europäischen Rat **koordinierte Vorschriften zur Förderung des transparenten grenzüberschreitenden Datenaustauschs**, die den Zugang zu umfangreichen Datenpools für Forschungs- und Innovationszwecke erleichtern.
3. **Förderung einer öffentlichen digitalen Infrastrukturgouvernanz** in den G20-Ländern, die sich auf digitale Identifizierungs- und Zahlungssysteme sowie auf Datenaustauschsysteme konzentriert.
4. **Open-Source-Codierung von KI-Technologien¹³**, was die Schaffung von KI-Tools für die Entwicklungszusammenarbeit und die humanitäre Hilfe ermöglicht und begrenzte technische Kenntnisse erfordert.

RISIKEN

1. **Fehlen einer Vereinbarung über internationale Cybersicherheitsregeln und -standards und/oder Nichteinhaltung der Weisung 321-0 des EDA über die Nutzung der elektronischen Infrastruktur, den Umgang mit Social Media und den Daten- und Informationsschutz im Web¹⁴:** Risiko für den Einsatz von KI für/gegen Sicherheits- und/oder Gesundheitsdaten und böswillige, schwer identifizierbare Angriffe auf sensible Systeme und Daten. Dabei stellt sich auch die Frage, wo oder unter welcher Gerichtsbarkeit die Daten physisch gehostet werden. Datenschutzvorgaben Schweiz: keine Eingabe von Personendaten – auch keine privaten Bilder, Verwendung von anonymisierten oder pseudonymisierten Daten, Einhaltung der Vorgaben zur Informationssicherheit: KEINE Eingabe von Daten, die als INTERN/VERTRAULICH/GEHEIM eingestuft sind, oder von nicht klassifizierten Daten, die andere sensible, von der öffentlichen Kommunikation ausgeschlossene Inhalte betreffen. Die Nutzung von KI-Tools im EDA braucht zuerst Bewilligung.
2. **Negative Auswirkungen durch verzerrte oder diskriminierende Daten:** Da KI-Systeme meist von grossen Privatunternehmen im Globalen Norden und in englischer Sprache entwickelt werden, enthalten sie in der Regel diskriminierende Verzerrungen, die in den Trainingsdaten oder -zielen aufgegriffen und verstärkt werden. Das führt zu einer Ungleichbehandlung mit negativen Folgen, insbesondere für Frauen und Mädchen, Minderheiten und den Globalen Süden, mit Auswirkungen auf Gouvernanzentscheide, soziale Inklusion und Menschenrechte.
3. **Grosse private Technologieunternehmen (meist aus dem Globalen Norden) dominieren die Systeme der Entwicklungszusammenarbeit und der humanitären Hilfe und sind sogenannte Gatekeeper.** Diese Unternehmen werden faktisch zu Dienstleistungsanbietern in der Entwicklungszusammenarbeit und der humanitären Hilfe, unterliegen aber nicht der gleichen Rechenschaftspflicht und Regulierungsaufsicht wie Hilfswerke. Es fehlt an Transparenz, Sicherheitsvorkehrungen und klaren Grenzen, wodurch diese Unternehmen in der Lage sind, eigenmächtig, d.h. ohne Sorgfaltsprüfung, Datenschutzprotokolle oder Mechanismen zur Vermeidung von Interessenkonflikten und regulatorischer Voreingenommenheit, neue KI-Technologien in volatilen Kontexten einzusetzen.
4. **Geopolitische Spannungen, die von autoritären Tendenzen und Modellen ausgehen**, die eine von individuellen Rechten, politischem Pluralismus und gerechtem Wohlstand abgekoppelte KI-Entwicklung verfolgen, die mit den Werten der Schweiz nicht vereinbar ist.
5. **Überwachung und Manipulation** in der Gouvernanz und der Bürgerbeteiligung sowie in Regionen mit Konflikten und Unruhen, wodurch vulnerable Bevölkerungsgruppen und Einzelpersonen gefährdet und Daten-/Digitalrechte und die Privatsphäre verletzt werden.
6. **Fehlinformation/Desinformation** (siehe Definitionen im Glossar), die demokratische und partizipatorische Prozesse beeinflussen und insbesondere die Geschlechtergleichheit und die Menschenrechte verletzen (z.B. indem Mädchen und Frauen gezielt von der Teilhabe ausgeschlossen oder gesellschaftliche Ansichten in einer Weise geprägt werden, die die Geschlechtergleichheit und die demokratische Inklusion behindert).
7. **Örtlich begrenzte und fragmentierte KI-Regulierung:** Die KI-Gouvernanz bleibt fragmentiert und örtlich begrenzt. Bestrebungen in der KI-Gouvernanz sind trotz gemeinsamer Grundsätze mit Unstimmigkeiten in Bezug auf die Umsetzung und die rechtsübergreifende Interoperabilität sowie mit unterschiedlichen Anreizen zum Einhalten von Vorschriften konfrontiert.
8. Die **KI-Lieferketten sind nur teilweise globalisiert** und durch internationale Vorgaben und Standards geregelt, weshalb sie vom Ansatz, den die Schweiz bei der Festlegung ihrer gesetzgeberischen Prioritäten im KI-Bereich verfolgt, abweichen und sich auf die Schweizer Wirtschaft, Forschung und humanitäre Praxis auswirken können.

4. 10 DEZA-Leitsätze für einen verantwortungsvollen KI-Einsatz

Auf der Grundlage der SWOT-Analyse, von Kapitel III des *Rahmenübereinkommens des Europarats über künstliche Intelligenz und Menschenrechte, Demokratie und Rechtsstaatlichkeit* sowie der OECD-Grundsätze für künstliche Intelligenz sollen die folgenden **10 Leitsätze für einen verantwortungsvollen KI-Einsatz bei der Arbeit der DEZA** als Geberin, Kooperationspartnerin oder Beraterin **systematisch berücksichtigt** und die DEZA/EDA-Mitarbeitenden geschult werden (vgl. Kapitel 8):

- 1) **Do no harm (Richte keinen Schaden an):** das Vorsorgeprinzip «Do no harm»¹⁵ anwenden, nichts übereilen¹⁶ und sich bei der Einführung von KI in die Prozesse an die Leitsätze halten: vgl. vorliegende DEZA-Arbeitshilfe und OECD-Grundsätze für KI¹⁷.
- 2) **Mensch behält Kontrolle**¹⁸: KI-Systeme auf eine beratende Funktion beschränken, z.B. zur Unterstützung von medizinischen Diagnosen, wobei die Beurteilung und Entscheidungsfindung weiterhin in menschlicher Hand bleibt. Andernfalls besteht die Gefahr von KI-Halluzinationen und Verzerrungen¹⁹.
- 3) **Beteiligung und Inklusion:** KI-Lösungen und **KI-Kenntnisse** gemeinsam mit und durch die betroffenen Bevölkerungsgruppen und Gemeinschaften entwickeln und dabei sicherstellen, dass in jeder Phase des Projektzyklus sowohl Frauen als auch Männer einbezogen werden. Vgl. *OECD Catalogue of Tools & Metrics for Trustworthy AI*²⁰.
- 4) **Lokalisierung von KI**²¹: sicherstellen, dass KI-Tools nicht nur von grossen ausländischen Unternehmen mit ihren jeweiligen Ideologien und Werten entwickelt werden, die den marginalisierten Gemeinschaften Macht und Ressourcen entziehen, sondern von Unternehmen, die der Realität, auf die das KI-Tool ausgerichtet sein soll, am nächsten sind, beispielsweise durch die Unterstützung regionaler KI-Hubs und die Förderung von Open-Source-Lösungen, die lokale Daten berücksichtigen.
- 5) **Bereinigung von Datensätzen (keine Verzerrungen):** Verzerrungen müssen während des gesamten KI-Lebenszyklus beachtet werden, von der Datenerfassung bis zum Einsatz des Modells: Verzerrungen in Datensätzen, Algorithmen und KI-gesteuerten Prozessen identifizieren und beheben. Es gibt verschiedene Methoden zur Bereinigung von Datensätzen²², z.B. zur Beseitigung historischer Muster von geschlechtsspezifischen und rassistischen Vorurteilen (z.B. verzerrte Daten, die geschlechtsspezifische Ungleichheiten aufgreifen, wodurch die KI bestehende Vorurteile reproduziert).
- 6) **Datenintegrität:** Für robuste, aussagekräftige KI-Anwendungen sind qualitativ hochwertige externe Datensätze erforderlich. Eine verantwortungsvolle Datennutzung setzt voraus, dass die Daten angemessen lizenziert sind und nach ethischen Gesichtspunkten beschafft werden; Datensätze unklarer/unbestimmter Herkunft oder die ohne Zustimmung bzw. rechtswidrig beschafft wurden, sollten nicht verwendet werden.
- 7) **Menschenwürdige Arbeit:** Ausbeutung in der KI-Lieferkette bekämpfen und faire Arbeitsbedingungen sicherstellen²³.
- 8) **Auswirkungen auf das Klima:** die Auswirkungen von KI-Technologien auf das Klima durch verantwortungsvolle Entscheidungen minimieren und/oder ausgleichen, z.B. energieeffiziente Modelle verwendet, die Rechenressourcen optimiert und wo immer möglich Cloud-Dienste bevorzugt werden, die über erneuerbare Energien laufen²⁴.
- 9) **Transparenz/algorithmisches Audit:** Wenn immer möglich die Codierung zur Verfügung stellen (Open Source). Für die Prozesse Kontrollen und eine Rechenschaftspflicht gewährleisten. Für das Daten-/KI-System eine algorithmische Rechenschaftspflicht gewährleisten²⁵ (siehe Glossar) und die KI-Systeme einer Menschenrechtsprüfung unterziehen²⁶.
- 10) Bewilligung durch das EDA: Der Einsatz eines KI-Tools innerhalb des EDA bedarf gemäss der EDA-Weisung 321-0 über die Nutzung der elektronischen Infrastruktur, den Umgang mit sozialen Medien und den Daten- und Informationsschutz im Web der Bewilligung durch das EDA²⁷. **Die DEZA als Geldgeberin im KI-Bereich (Finanzierung von operationellen Projekten)**

Alle Abteilungen und Sektionen der DEZA sowie alle Länder setzen sich vermehrt mit KI-Technologien auseinander. Jede DEZA-Einheit kann im Rahmen ihres Budgets und gemäss den üblichen Regeln für Eintretens- oder Kreditanträge ein Projekt mit KI-Komponenten einreichen, sofern die vorliegende Arbeitshilfe und insbesondere die Leitsätze für einen verantwortungsvollen KI-Einsatz eingehalten werden.

Die DEZA verfolgt bei der Finanzierung von operativen KI-Projekten einen dreifachen Ansatz, oft noch mit einem experimentellen Teil, manchmal auch mit einer Kombination von Ansätzen.

- a) **Konkrete direkte Projekte mit «sichererer» operativer Anwendung von KI:** Für Projekte mit geringem Reputationsrisiko und geringer politischer Sensibilität eignet sich eine Technologie, die auf wissenschaftlich gesammelten, nicht sensiblen Daten beruht (z.B. Meteorologie, Seismologie, Satellitendaten, Logistik oder Inzidenz von Krankheiten). Die DEZA verfügt über einige Erfahrung in den Bereichen *Klima* in Laos und Indonesien (Risiken im Zusammenhang mit El Niño²⁸), *Gesundheit* in Tansania und Ruanda (pädiatrisches Diagnose-Tool mit dem Ziel, die Kindersterblichkeit zu senken und unnötige Antibiotikaverschreibungen zu vermeiden²⁹) sowie *Global* (Zentrale, Projekt «Governing Health Futures 2030 Youth Network»³⁰) und Kernbeiträge an Schweizer NGOs (Stiftung Hiron-delle³¹).
- b) **Mittel- bis langfristige Projekte, die sich auf KI-Bedürfnisse und den Aufbau lokaler und regionaler KI-Kapazitäten im Globalen Süden und in den Partnerländern konzentrieren:** Dies entspricht insbesondere Punkt 3 (Partizipation und Inklusion), Punkt 4 (Lokalisierung von KI) und Punkt 5 (Keine Verzerrung von Datensätzen) der oben aufgeführten *Leitsätze für einen verantwortungsvollen KI-Einsatz* und vor allem dem Mandat der DEZA als Entwicklungsagentur mit lokal geführter Zusammenarbeit. Zu erwähnen sind hier das **Schweizer ICAIN-Projekt**³² oder das Afrika-Projekt AI4D – gemeinsam finanziert von Kanada/IDRC, UK/FCDO, Schweden/SIDA, Deutschland/GIZ und der Gates Stiftung³³.
- c) **Konkrete gemeinsame Projekte, die mit anderen Gebern kofinanziert werden: für Projekte, bei denen es sicherer ist, die Risiken zu teilen, weil sie politisch sensibel sind (z.B. Frieden, Gouvernanz und Gender), und/oder um die Wirkung über die Reichweite der DEZA hinaus zu maximieren.** Gemäss dem ersten Leitsatz (Do no harm) muss die DEZA-Vorkehrungen gegen Reputationsrisiken und bekannte negative Auswirkungen von KI auf die Gesellschaft treffen. Die DEZA soll nicht mit operativen KI-Projekten experimentieren, die verletzte Bevölkerungsgruppen betreffen oder biometrische und personenbezogene Daten verwenden. Hier besteht ein hohes Risiko für negative Auswirkungen – bestehende geschlechts- und rassenspezifische Ungleichheiten könnten verstärkt werden. Das Staatssekretariat für Wirtschaft (SECO) hat seinerseits die GovTech Global Partnership³⁴ finanziert. **Ein Überblick über ausgewählte KI-Projekte wichtiger Geber, die dieselben KI-Grundsätze wie die DEZA verfolgen, ist verfügbar**³⁵.

6. Die DEZA als Kooperationspartnerin im KI-Bereich (Einflussnahme auf die Politik)

Durch ihre Mitgliedschaft im DAC/OECD, ihren regelmässigen (Vize-)Vorsitz in einigen DAC-Netzwerken sowie durch die Schweizer Vertretungen in New York, Genf und Paris kann und soll die DEZA eine wichtige Rolle bei der **politischen Einflussnahme auf das Thema KI im Bereich der internationalen Zusammenarbeit** spielen.

Über das **Staatssekretariat** ist das EDA auf diplomatischer Ebene aktiv im KI-Bereich, mit einigen konkreten Initiativen, die z.B. den politischen Dialog der DEZA unterstützen können:

- **Geneva Digital Atlas**, einschliesslich eines speziellen Chatbots³⁶: umfassende Bestandsaufnahme der Akteure in den Bereichen Digitalpolitik und Internet-Gouvernanz im internationalen Genf. Der Atlas enthält eine ausführliche Auflistung der Aktivitäten von über 40 Akteuren.
- **Geneva Science and Diplomacy Anticipator (GESDA)**, mit seinem Open Quantum Institute, das sich für die Förderung offener Daten für eine nachhaltige Entwicklung einsetzt³⁷.
- **International Computation and AI Network (ICAIN)**³⁸, in Zusammenarbeit mit der ETH Zürich, *Data Science Africa* und weiteren internationalen Partnern, darunter zwei Hochleistungsrechenzentren – verfolgt das Ziel, KI-Technologien mit gesellschaftlichem Mehrwert zu entwickeln. Zudem soll Akteurinnen und Akteuren in Ländern mit niedrigem und mittlerem Einkommen der Zugang zu Supercomputern, Dateninfrastrukturen und KI-Expertise ermöglicht werden, um KI-basierte Lösungen für das Gemeinwohl zu fördern. In der Anfangsphase konzentrierte sich das Projekt auf Landwirtschaft und Klimawandel, Gesundheit und Humanitäres (Meditron und *Large Language Model* des IKRK [siehe Definitionen im Glossar]). Die Humanitäre Hilfe der DEZA ist federführend bei der Entwicklung weiterer Projekte, die von ICAIN profitieren.
- **Cybermediation** mit der Diplo Foundation³⁹.

Neben den genannten konkreten diplomatischen Initiativen führt das Staatssekretariat **jährlich bilaterale Gespräche speziell zum Thema KI mit anderen Ländern**, beispielsweise mit dem Vereinigten Königreich. Die DEZA kann solche Gespräche für den Aufbau von Partnerschaften nutzen. Schliesslich hat das Staatssekretariat den Vorsitz in der interdepartementalen Arbeitsgruppe *Strategie Digitalausserpolitik* (einschliesslich KI), mit der DEZA als Mitglied.

Die DEZA und das Staatssekretariat EDA setzen sich gestützt auf die rechtebasierten Werte und die wirtschaftspolitische Stärke der Schweiz **gemeinsam für eine verantwortungsvolle internationale Politik im KI-Bereich** ein. Viele der Risiken und negativen Auswirkungen der KI betreffen Menschenrechtsverletzungen und verstärken bestehende Wirtschafts- und Machtinteressen sowie Ungleichheiten. Der Europarat, zu dessen Gründungsmitgliedern auch die Schweiz zählt, und das Europäische Parlament haben **Empfehlungen zu KI und Menschenrechten**⁴⁰ veröffentlicht, die im Mittelpunkt der politischen Debatte und der Einflussnahme der Schweiz stehen sollten. Auf bilateraler Ebene spielt die **Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ)** eine führende Rolle beim Thema KI für die Entwicklungszusammenarbeit und die humanitäre Hilfe und den Regeln in diesem Bereich. Die GIZ ist federführend bei der Entwicklung einer **Bewertungsmethode für verantwortungsvolle KI**⁴¹ und hat ein **Handbuch für politische KI-Entscheidungsträger**⁴² veröffentlicht.

Die wichtigsten internationalen Veranstaltungen und Foren für die neusten Informationen und den Politikdialog über KI sind die folgenden (die ersten vier sind für die DEZA am wichtigsten):

- **Internationaler «AI for Good»-Gipfel**⁴³, jährlich, in Genf: In Zusammenarbeit mit der Internationalen Fernmeldeunion (ITU). Fokus auf KI-Richtlinien und -Instrumente für die humanitäre Hilfe und die internationale Zusammenarbeit, mitorganisiert von EDA und BAKOM.
- **Konferenz «AI for Development (AI4D)»**⁴⁴, jährlich, in Barcelona (Spanien): Fokus auf praktische KI-Anwendungen in der internationalen Zusammenarbeit, um Herausforderungen wie Nachhaltigkeit und Ethik anzugehen.

- **Global AI Summit on Africa**⁴⁵: Fokus auf Afrika. Mitorganisiert vom Weltwirtschaftsforum.
- **Weltgipfel zur Informationsgesellschaft (WSIS)**⁴⁶, in Genf: Fokus nicht nur auf KI, sondern im weiteren Sinne auf Kommunikationstechnologien sowie auf das Problem von Des-, Falsch- und Fehlinformation (siehe Definitionen im Glossar). Im Rahmen der Internationalen Fernmeldeunion (ITU).
- **KI-Gipfeltreffen, die regelmässig von verschiedenen Ländern veranstaltet werden**: Fokus auf KI im Allgemeinen für alle Bereiche der Gesellschaft, also nicht nur für die internationale Zusammenarbeit, wie der KI-Gipfel von 2025 in Paris⁴⁷, möglicherweise später in der Schweiz.
- **Internet Governance Forum (IGF) der UNO**⁴⁸, jährlich: vereint Menschen aus verschiedenen Interessengruppen zu einer Debatte über die digitale öffentliche Politik und ist daher nicht ausschliesslich dem Thema KI gewidmet. Wird jeweils von einem anderen Land ausgerichtet.
- **Hamburg Sustainability Conference (HSC)**⁴⁹, jährlich, in Hamburg (Deutschland): Fokus auf Nachhaltigkeit der Entwicklung im Allgemeinen, einschliesslich einer Plattform über das Potenzial verantwortungsvoller KI für die nachhaltige Entwicklung mit einem «AI for SDGs Compendium», einem globalen Verzeichnis von KI-Initiativen für nachhaltige Entwicklung⁵⁰.

Policy Messages der DEZA zum Thema KI für die politische Einflussnahme und den Dialog auf allen Ebenen:

Aktuelle globale Debatte	Policy Messages	Potenzielle Partner
<ul style="list-style-type: none"> • «AI4 Good»-Gipfel (Genf) • AI4D-Konferenz • Global AI Summit on Africa: • Weltgipfel zur Informationsgesellschaft (WSIS) • KI-Gipfeltreffen (Rotation) • Internet Governance Forum (IGF) der UNO • Hamburg Sustainability Conference (HSC) 	<ul style="list-style-type: none"> - Die DEZA ist sich bewusst, dass die KI die Zukunft der Entwicklungszusammenarbeit und der humanitären Hilfe prägt – im Guten wie im Schlechten – und setzt sich aktiv für eine verantwortungsvolle, regulierte und inklusive KI ein, welche die digitale Spaltung nicht weiter verschärft. - Die DEZA unterstützt Grundsätze für eine verantwortungsvolle KI mit Blick auf Anwendungen in der humanitären Hilfe und der Entwicklungszusammenarbeit (mögliche Partnerschaft mit AI4D undCAIN). - Die DEZA trägt zur Formulierung von KI-Richtlinien, -Normen und -Werten bei, die negative Auswirkung minimieren, die soziale Inklusion fördern, die Geschlechtergleichheit stärken und mit den Menschenrechtsstandards in Einklang stehen. - Die DEZA ist bestrebt, lokale Talente und Fähigkeiten im KI-Bereich aufzubauen, und unterstützt regionale KI-Hubs, die auf lokale Fähigkeiten und eine verantwortungsvolle KI-Politik ausgerichtet sind. - Die DEZA trägt zur Förderung der inklusiven KI bei, indem sie die Entwicklung lokaler Open-Language-Datensätze unterstützt. 	<ul style="list-style-type: none"> • CIDA-IDRC/UK-FCDO/Schweden-SIDA/Deutschland/US Gates Foundation (AI4D) • Eidgenössische Technische Hochschulen (ICAIN) • FAIR Forward (GIZ) • Global Index on Responsible AI • Digital Impact Alliance • D4D-Hub (EU) • ITU

7. Die DEZA als Beraterin über KI (thematische Beratung)

Die thematische Beratung der DEZA (thematische Sektionen und thematische Mitarbeitende in den geografischen Sektionen) verfolgt für alle DEZA-Themen folgende Ziele:

- Sicherstellen, dass die vorliegende Arbeitshilfe und die im EDA geltenden Rechtsgrundlagen bekannt sind und angewendet werden. Bei Bedarf werden diese Instrumente vorgestellt;
- Die SWOT-Analyse zu KI für jedes Thema und jede Region bzw. jedes Land vertiefen;
- Projekte vorschlagen, die von verschiedenen DEZA-Einheiten finanziert werden;
- In allen Themenbereichen der DEZA zu KI-Aspekten des politischen Dialogs und der Einflussnahme beraten;
- Den antragstellenden Einheiten Orientierungshilfen zur Verfügung stellen, mit denen sie ermitteln können, ob KI-Projekte in einem bestimmten Themenbereich die Menschenrechte wahren, die Geschlechtergleichheit fördern und mit den Leitsätzen für einen verantwortungsvollen KI-Einsatz in Einklang stehen;
- Die antragstellenden Einheiten bei ihren eigenen KI-Überlegungen kontextspezifisch unterstützen, insbesondere im Hinblick auf Gefahren, Prioritäten und Innovationsmöglichkeiten, im Einklang mit der vorliegenden KI-Arbeitshilfe und den thematischen DEZA-Richtlinien.

Auch hier müssen die DEZA/EDA-Mitarbeitenden und ihre Partner spezifische KI-Kapazitäten aufbauen, insbesondere durch spezielle Schulungen (siehe Kapitel 8) und mit Unterstützung von spezialisierten externen Backstoppers.

8. Die DEZA als Wissensvermittlerin durch KI (Wissensmanagement)

Mittel- bis langfristig ist zu erwarten, dass die meisten Tätigkeiten der DEZA und die von ihr unterstützten Projekte konkrete KI-Aspekte beinhalten. Damit die DEZA unter Berücksichtigung dieser Arbeitshilfe bestmöglich für die Zukunft gewappnet ist, gelten bis auf Weiteres folgende Massnahmen:

- **Erfassung/Monitoring:** Projekte, die KI-Technologien beinhalten, müssen mit der Bezeichnung «*künstliche Intelligenz (KI)*» gekennzeichnet sein (ausgeschrieben), insbesondere im Projekt-Factsheet des Eintretens- oder Kreditantrags. Auf diese Weise kann in der offiziellen, öffentlich und transparent zugänglichen DEZA-Projektdatenbank direkt nach KI-Projekten gesucht und gefiltert werden, ohne dass eine separate Erfassung erforderlich ist. Intern und im Rahmen des OECD-DAC wird aktuell mit Blick auf eine Entscheidung darüber debattiert, ob für DEZA-Beiträge an KI-Projekte in SAP ein besonderes Monitoring wünschenswert oder notwendig ist, entweder mit dem derzeitigen Marker für Digitalisierung in Kombination mit Stichwörtern oder mit einem eigenen KI-(Sub-)Marker.
- **Ausbildung:**
 - Die Bundesverwaltung bietet über die Plattform SuccessFactors⁵¹ eine **Grundausbildung in Datenkompetenz, Datenmanagement, Datenanalyse, Informationssicherheit und Einführung in die KI** an. **DEZA-Mitarbeitende müssen diese Kurseinheiten absolvieren.** Die anderen Kurse zum Thema KI richten sich speziell an IT-Entwicklerinnen und -Entwickler und sind für die meisten DEZA-Mitarbeitenden nicht geeignet.
 - Ausserhalb der Bundesverwaltung gibt es **spezielle KI-Schulungen für Fachleute aus den Bereichen Entwicklungszusammenarbeit und/oder humanitäre Hilfe. Eine Liste ist online abrufbar** und wird regelmässig aktualisiert⁵². Die Linienvorgesetzten sorgen dafür, dass Mitarbeitende, die unmittelbar mit KI-Technologien zu tun haben, gewisse Kurse besuchen können. Die Mitarbeitenden ihrerseits berichten der oder dem Vorgesetzten über die absolvierten Kurse.
- **Austausch:** Die DEZA berichtet der interdepartementalen Arbeitsgruppe *Strategie Digitalaussenpolitik* regelmässig über die Erfahrungen der DEZA mit KI sowie über den Austausch zwischen Gebern und Partnern im KI-Bereich (z.B. über OECD-Netzwerke) und umgekehrt. Die betroffenen Abteilungen und Sektionen werden entsprechend informiert.

9. KI in DEZA-internen Prozessen

Potenziell können KI-Technologien das **Projektzyklusmanagement und die Qualitätssicherung** der DEZA erheblich stärken, indem durch prädiktive KI, generative KI und KI-basiertes Content Monitoring (vgl. Glossar) **Routineaufgaben automatisiert, die Datenanalyse ausgeweitet und die Entscheidungsprozesse verbessert** werden. Die DEZA-interne Nutzung von KI wirft weniger rechtliche und ethische Fragen auf. Die Analyse organisationsinterner Daten durch maschinelles Lernen ist weit weniger problematisch als das Kopieren grosser Datenmengen aus dem Internet.

Die DEZA hat bereits im Jahr 2024 konkrete Schritte zur Integration von KI in interne Prozesse unternommen:

- Ein erfolgreiches Pilotprojekt zur automatischen Zuordnung von Projektmerkmalen (Technologietransfer und Kapazitätsaufbau) auf der Grundlage der Projektbeschreibungen. Die Ergebnisse wurden für die Berichterstattung an das UNFCCC verwendet.
- KI wird in laufenden Digitalisierungsinitiativen wie *EDaTa* und *SDC Projects* berücksichtigt.

Um die Vorteile zu maximieren, sollte die DEZA auch in Zukunft **KI-Tools bevorzugen, die sich in bestehende Arbeitsabläufe integrieren lassen** und gleichzeitig den Datenschutz, die Informationssicherheit, ethische Erwägungen und die notwendige Schulung der Mitarbeitenden gewährleisten. Die Ausbildung sollte nicht nur KI-Kenntnisse vermitteln, sondern auch die bestehende IT-Architektur der Bundesverwaltung und die geltenden rechtlichen Anforderungen erläutern, und zwar in enger Zusammenarbeit mit IT-EDA, um eine nachhaltige und sichere Umsetzung sicherzustellen.

Eine Auflistung möglicher Anwendungen von KI im Rahmen von Programmen der Entwicklungszusammenarbeit und der humanitären Hilfe **für einzelne PCM-Schritte und Themen oder für die methodische Qualitätssicherung ist verfügbar** und wird regelmässig aktualisiert. Die DEZA prüft die bereits bestehenden Möglichkeiten und kofinanziert gegebenenfalls besonders geeignete Lösungen⁵³. Gewisse **Apps in Microsoft Teams** können in der DEZA ebenfalls genutzt werden, sofern sie von IT-EDA freigegeben sind.

Glossar

Algorithmisches Audit: Evaluation, Milderung und Gewährleistung der Sicherheit, Rechtmässigkeit und Ethik von Algorithmen und diesbezügliche Forschung.

Desinformation: Informationen, die falsch sind und von denen die Person, die sie erstellt oder verbreitet, weiss, dass sie falsch sind.

Falschinformation: Informationen, die auf der Realität beruhen, aber dazu verwendet werden, einer Person, einer Organisation oder einem Land zu schaden.

Fehlinformation: Informationen, die zwar falsch sind, aber nicht in der Absicht der Irreführung erstellt oder verbreitet wurden und von denen die Person, die sie verbreitet, glaubt, dass sie wahr sind.

Generative KI: KI, die neue Inhalte wie Bilder, Musik oder Text erstellt, indem sie aus vorhandenen Beispielen lernt. Sie funktioniert als kreative Maschine, die Dinge völlig neu erschaffen kann.

KI-basiertes Content Monitoring: Einsatz von KI-Technologien zur automatischen Überprüfung, Analyse und Verwaltung digitaler Inhalte auf verschiedenen Plattformen, um sicherzustellen, dass diese mit vordefinierten Standards, Leitlinien oder Richtlinien übereinstimmen. Dazu gehören die Erkennung und Verwaltung unangemessener, schädlicher oder unzulässiger Inhalte in Text, Bild, Video und Audio.

Künstliche Intelligenz (KI): Fähigkeit von Computersystemen oder Maschinen, Aufgaben auszuführen, die normalerweise menschliche Intelligenz erfordern, wie das Erkennen von Gesichtern, das Treffen von Entscheidungen oder das Verstehen von Sprache.

Large Language Model (LLM): Art von KI, die menschliche Sprache verstehen und erzeugen kann. Sie kann Fragen beantworten, Aufsätze verfassen oder sich mit Menschen unterhalten, indem sie riesige Mengen von Textdaten verarbeitet.

Maschinelles Lernen: Methode, mit der Computern beigebracht wird, aus Erfahrung zu lernen, so wie der Mensch aus der Praxis lernt. Computer können dadurch ihre Leistung bei der Ausführung von Aufgaben verbessern, ohne dass sie explizit für jede einzelne Aufgabe programmiert werden müssen.

Prädiktive KI: KI, die Daten und Muster nutzt, um künftige Ereignisse zu prognostizieren. So kann sie beispielsweise das Wetter oder das Kaufverhalten von Kundinnen und Kunden vorhersagen.

Verantwortungsvolle KI: Gestaltung, Entwicklung und Einsatz von KI-Systemen, die sicher, inklusiv, rechtebasiert und nachhaltig sind.

-
- ¹ Genaue Linkadresse: https://static.vecteezy.com/system/resources/previews/022/190/659/non_2x/double-exposure-image-of-virtual-human-3dillustration-on-blue-circuit-board-background-represent-artificial-intelligence-ai-technology-free-photo.jpg.
- ² *Rahmenübereinkommen des Europarats über künstliche Intelligenz und Menschenrechte, Demokratie und Rechtsstaatlichkeit* (05.09.2024), nur auf English oder Französisch verfügbar: <https://www.coe.int/en/web/conventions/full-list?module=treaty-detail&treatyid=225>.
- ³ *Bundesgesetz über den Datenschutz*: <https://www.fedlex.admin.ch/eli/cc/2022/491/de>.
- ⁴ *EU-Datenschutz-Grundverordnung*: <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A32016R0679>.
- ⁵ *OECD-Grundsätzen für künstliche Intelligenz*, nur auf English oder Französisch verfügbar: <https://oecd.ai/en/ai-principles>.
- ⁶ *UNESCO-Empfehlung zur Ethik der künstlichen Intelligenz*, nur auf English oder Französisch verfügbar: <https://unesdoc.unesco.org/ark:/48223/pf0000380455>.
- ⁷ *[Schweizer] Strategie zur internationalen Zusammenarbeit (IZA) 2025–28*: <https://www.eda.admin.ch/eda/de/home/das-eda/aktuell/dossiers/iza-strategie-2025-28.html>.
- ⁸ [Schweizer] Eidgenössisches Departement des Innern, Bundesamt für Statistik, Kompetenznetzwerk für künstliche Intelligenz (CNAI), *Merkblätter zu KI in der Bundesverwaltung*: <https://cna1.swiss/>.
- ⁹ *Bundesgesetz über die internationale Entwicklungszusammenarbeit und humanitäre Hilfe*: https://www.fedlex.admin.ch/eli/cc/1977/1352_1352_1352/de.
- ¹⁰ Swiss National Supercomputing Centre, nur auf English verfügbar: <https://www.cscs.ch/>.
- ¹¹ Project *International Computation and AI Network (ICAIn)*, nur auf English verfügbar: <https://icain.ch/>.
- ¹² EU KI-Gesetz: <https://digital-strategy.ec.europa.eu/de/policies/regulatory-framework-ai>.
- ¹³ Forbes, Bernard Marr, *The 10 Best Examples Of Low-Code And No-Code AI* (12.12.2022), nur auf English verfügbar: <https://www.forbes.com/sites/bernardmarr/2022/12/12/the-10-best-examples-of-low-code-and-no-code-ai/>.
- ¹⁴ *Weisung 321-0 des EDA über die Nutzung der elektronischen Infrastruktur, den Umgang mit Social Media und den Daten- und Informationsschutz im Web* (mehrsprachig): <https://intranet.eda.admin.ch/en/publication?id=Kqj1wQXQyWws>.
- ¹⁵ ETUI, Ponce del Castillo A. (2024) *AI: the value of precaution and the need for human control*, in Ponce del Castillo A. (ed.) *Artificial intelligence, labour and society*, ETUI, nur auf English verfügbar: https://www.etui.org/sites/default/files/2024-03/Chapter1_AI%20the%20value%20of%20precaution%20and%20the%20need%20for%20human%20control.pdf.
- ¹⁶ Fu Gratiana, *Case Studies in the Practice of Responsible AI for Development*. Farnham, Surrey, UK: Caribou Digital Publishing, September 2024, nur auf English verfügbar: <https://www.cariboudigital.net/publication/case-studies-in-responsible-ai-for-development>.
- ¹⁷ *OECD-Grundsätzen für künstliche Intelligenz*, nur auf English oder Französisch verfügbar: <https://oecd.ai/en/ai-principles>.
- ¹⁸ Nature, npj digital medicine, *How AI can learn from the law: putting humans in the loop only on appeal* (2023), nur auf English verfügbar: <https://www.nature.com/articles/s41746-023-00906-8>.
- ¹⁹ MIT Sloan Teaching & Learning Technologies, *When AI Gets It Wrong: Addressing AI Hallucinations and Bias* (2025), nur auf English verfügbar: <https://mitsloanedtech.mit.edu/ai/basics/addressing-ai-hallucinations-and-bias/>.
- ²⁰ OECD.AI Policy Observatory and GPAI, *Catalogue of Tools & Metrics for Trustworthy AI* (27.03.2023), nur auf English oder Französisch verfügbar: <https://oecd.ai/en/catalogue/tools/participatory-ai-framework>.
- ²¹ Stanford University, Human-Centered Artificial Intelligence, *The Movement to Decolonize AI: Centering Dignity Over Dependency* (21.03.2022), nur auf English verfügbar: <https://hai.stanford.edu/news/movement-decolonize-ai-centering-dignity-over-dependency>.
- Chatham House, *Why AI must be decolonized to fulfill its true potential* (10-11.2023), nur auf English verfügbar: <https://www.chathamhouse.org/publications/the-world-today/2023-10/why-ai-must-be-decolonized-fulfill-its-true-potential>.
- ²² MIT News, Adam Zewe, 11.12.2024, nur auf English verfügbar: <https://news.mit.edu/2024/researchers-reduce-bias-ai-models-while-preserving-improving-accuracy-1211>.
- ²³ Fairwork, *Die tägliche Arbeitsrealität verändert sich*, <https://fair.work/de/fw/about/>.
- ²⁴ Global Partnership on Artificial Intelligence (GPAI), *A responsible AI strategy for the environment*, nur auf English verfügbar: <https://gpai.ai/projects/responsible-ai/environment/>.
- ²⁵ Santa Clara High Technology Law Journal, Volume 39 Issue 3 Article 1, Goodman Ellen P. & Trehu Julia, *Algorithmic Auditing: Chasing AI Accountability* (26.05.2023), nur auf English verfügbar: https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?params=/context/chtj/article/1689/&path_info=11_Goodman_Algorithmic_Auditing_Chasing_AI_Accountability_PUBLISHED.pdf.
- ²⁶ Council of Europe, David Leslie, Christopher Burr, Mhairi Aitken, Josh Cowsls, Mike Katell & Morgan Briggs, *Artificial Intelligence, Human Rights, Democracy and the Rule of Law – Prepared to support the Feasibility Study published by the Council of Europe’s ad hoc Committee on Artificial Intelligence*, nur auf English verfügbar: https://www.turing.ac.uk/sites/default/files/2021-03/cahai_feasibility_study_primer_final.pdf.
- ²⁷ Kompetenzer Dienst: EDA Datenschutz, Informationssicherheit und Informationszugang, informationssicherheit@eda.admin.ch, <https://intranet.eda.admin.ch/de/informatikssicherheit>.

- ²⁸ El Niño-related Risks nur auf English verfügbar: <https://www.fdfa.admin.ch/deza/en/home/projekte/projekte.filterResults.html/content/dezaprojects/SDC/en/2024/7F11321/phase1?oldPagePath=/content/deza/en/home/projekte/projekte.html>.
- ²⁹ Medical diagnosis tool nur auf English verfügbar: <https://www.fdfa.admin.ch/deza/en/home/projekte/projekte.filterResults.html/content/dezaprojects/SDC/en/2019/7F10361/phase1?oldPagePath=/content/deza/en/home/projekte/projekte.html>.
- ³⁰ Global Governing Health Futures 2030 nur auf English verfügbar: <https://www.fdfa.admin.ch/deza/en/home/projekte/projekte.filterResults.html/content/dezaprojects/SDC/en/2020/7F10616/phase1?oldPagePath=/content/deza/en/home/projekte/projekte.html>.
- ³¹ Kernbeitrag an Stiftung Hironnelle, 2025-2026 nur auf English verfügbar: <https://www.fdfa.admin.ch/deza/en/home/projekte/projekte.html/content/dezaprojects/SDC/en/2013/7F08541/phase7>.
- ³² International Computation and AI Network (ICAIN), nur auf English verfügbar: <https://icain.ch/>.
- ³³ AI4D, nur auf English oder Französisch verfügbar: <https://www.ai4d.ai/>.
- ³⁴ GovTech Global Partnership nur auf English verfügbar: <https://www.fdfa.admin.ch/deza/en/home/projekte/projekte.filterResults.html/content/dezaprojects/SECO/en/2020/UR01266/phase1?oldPagePath=/content/deza/en/home/projekte/projekte.html>.
- ³⁵ **List of Major Donors' AI Projects** nur auf English verfügbar: <https://fafwpgovch.sharepoint.com/:f:/r/sites/DigitalTransformationSDC-AI/Shared%20Documents/SDC%20AI%20Working%20Aid?csf=1&web=1&e=2MKlJk>. Oder (für die Mitglieder des DEZA-Netzwerkes Frieden, Gouvernanz und Gleichstellung): https://fafwpgovch.sharepoint.com/:f:/r/sites/EDA_AGR_SDC-Networks-PGE-02.PGEMembers-Internal/Shared%20Documents/01.%20PGE%20Members-Internal%20only?csf=1&web=1&e=tRiBlQ.
- ³⁶ Geneva Digital Atlas, nur auf English verfügbar: <https://dig.watch/atlas>.
- ³⁷ GESDA, nur auf English verfügbar: <https://gesda.global/>.
- ³⁸ International Computation and AI Network (ICAIN) , nur auf English verfügbar: <https://icain.ch/>.
- ³⁹ Cybermediation, nur auf English verfügbar: <https://www.diplomacy.edu/topics/cybermediation/>.
- ⁴⁰ Council of Europe, May 2019, *Unboxing Artificial Intelligence - 10 steps to protect Human Rights – Recommendation*, nur auf English verfügbar: <https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64>.
European Parliament, H. Akin ÜNVER, May 2024: *Artificial intelligence (AI) and human rights: Using AI as a weapon of repression and its impact on human rights*, nur auf English verfügbar: [https://www.europarl.europa.eu/Reg-DATA/etudes/IDAN/2024/754450/EXPO_IDA\(2024\)754450_EN.pdf](https://www.europarl.europa.eu/Reg-DATA/etudes/IDAN/2024/754450/EXPO_IDA(2024)754450_EN.pdf).
- ⁴¹ Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) - FAIR Forward, *Responsible AI Assessments. Part A: Step-by-Step Guide*. Licensed under CC BY-SA 4.0 DEED., 2024, nur auf English verfügbar: <https://openforgood.info/wp-content/uploads/2024/11/Responsible-AI-Assessments-Part-A-Step-by-Step-Guide.pdf>.
- ⁴² Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) - FAIR Forward, *Handbook for Artificial Intelligence (AI) Training for Policymakers*, 2024, nur auf English verfügbar: <https://www.bmz-digital.global/wp-content/uploads/2024/09/ENG-highres-AI-POLICYMAKERS-TRAINING-HANDBOOK.pdf>.
- ⁴³ AI for Good Global Summit, nur auf English oder Französisch verfügbar: <https://aiforgood.itu.int/summit24/>.
- ⁴⁴ AI4D Annual Conference, nur auf English verfügbar: <https://www.ai4development.ai/>.
- ⁴⁵ Global AI Summit on Africa, nur auf English verfügbar: <https://c4ir.rw/global-ai-summit-on-africa/>.
- ⁴⁶ World Summit on the Information Society (WSIS) , nur auf English, Französisch oder Spanisch verfügbar: <https://www.itu.int/net/wsis/>.
- ⁴⁷ Paris AI Action Summit 2025, nur auf Französisch oder English verfügbar: <https://www.elysee.fr/en/sommet-pour-l-action-sur-l-ia>.
- ⁴⁸ UN Internet Governance Forum (IGF), nur auf English verfügbar: <https://www.intgovforum.org/en>.
- ⁴⁹ Hamburg Sustainability Conference, nur auf English verfügbar: <https://www.sustainability-conference.org/en/>.
- ⁵⁰ AI for SDGs Compendium, nur auf English verfügbar: <https://www.sdgaicompndium.org/principles>.
- ⁵¹ Platform SuccessFactors: https://hcm74.sapsf.eu/sf/home?bplte_company=bundesamtff.
- ⁵² **List of AI External Trainings**, nur auf English verfügbar: <https://fafwpgovch.sharepoint.com/:f:/r/sites/DigitalTransformationSDC-AI/Shared%20Documents/SDC%20AI%20Working%20Aid?csf=1&web=1&e=2MKlJk>. Oder (für die Mitglieder des DEZA-Netzwerkes Frieden, Gouvernanz und Gleichstellung): https://fafwpgovch.sharepoint.com/:f:/r/sites/EDA_AGR_SDC-Networks-PGE-02.PGEMembers-Internal/Shared%20Documents/01.%20PGE%20Members-Internal%20only?csf=1&web=1&e=tRiBlQ.
- ⁵³ **List of AI Technologies in PCM**, nur auf English verfügbar: <https://fafwpgovch.sharepoint.com/:f:/r/sites/DigitalTransformationSDC-AI/Shared%20Documents/SDC%20AI%20Working%20Aid?csf=1&web=1&e=2MKlJk>. Oder (für die Mitglieder des DEZA-Netzwerkes Frieden, Gouvernanz und Gleichstellung): https://fafwpgovch.sharepoint.com/:f:/r/sites/EDA_AGR_SDC-Networks-PGE-02.PGEMembers-Internal/Shared%20Documents/01.%20PGE%20Members-Internal%20only?csf=1&web=1&e=tRiBlQ.