
ENKIDU

— MIOVSKI GROUP —

Daten-Governance und -Ethik — Bewertungs- und Akkreditierungskriterien

Für maschinelles Lernen, neuronale Netze und KI-Systeme

Version 2.0 · Aktualisiert nach der Kriterienzuordnung

24. Januar 2026

Erstellt von MIOVSKI GROUP

Einstufung · Vertraulich · Entwurf

Zusammenfassung

Dieses Dokument legt umfassende Kriterien zur Bewertung und Akkreditierung von Praktiken der Daten-Governance und -Ethik in Systemen der künstlichen Intelligenz, des maschinellen Lernens und neuronaler Netze fest. Diese Fassung wurde anhand der offiziellen Kriterienzuordnung aktualisiert, die festlegt, welche Kriterien für Trainingsdaten, Inferenzdaten oder beide gelten.

Das Rahmenwerk gliedert sich in drei Hauptabschnitte:

- **ABSCHNITT 1: KRITERIEN DES DATENLEBENSZYKLUS** (Stufen 1 bis 10) — für Trainings- und Inferenzdaten bewertet, mit jeweils angegebener Anwendung.
- **ABSCHNITT 2: KRITERIEN DES DATENBETRIEBS** — in getrennte Unterabschnitte Training und Inferenz aufgeteilt, mit identischen Codes, aber unterschiedlicher Anwendung.
- **ABSCHNITT 3: ÜBERGREIFENDE GOVERNANCE- UND ETHIKKRITERIEN** — Kriterien auf Organisationsebene für alle KI-Systeme, nicht projektspezifisch.

Gesamtzahl der Bewertungskriterien: 63 eindeutige Kriterien über alle Kategorien.

Verteilung der Kriterien

- Nur Trainingsdaten: 5 Kriterien (Datenbetrieb — Training)
- Nur Inferenzdaten: 10 Kriterien (5 Betrieb + 5 Weitergabe)
- Sowohl Training als auch Inferenz: 36 Kriterien (von der Erfassung bis zur Entsorgung, ohne Betrieb)
- Übergreifend organisatorisch: 12 Kriterien (Governance, Dokumentation, Ethik, Schulung)

1. Einführung und Grundlagen des Rahmenwerks

1.1 Zweck und Geltungsbereich

Dieses Rahmenwerk bietet standardisierte, evidenzbasierte Kriterien zur Bewertung der Daten-Governance und der ethischen Konformität von KI-/ML-Systemen. Es gilt für Organisationen, die KI-Systeme entwerfen, entwickeln, bereitstellen oder betreiben, mit besonderem Fokus auf Anwendungen mit hoher Auswirkung und hohem Risiko, die Grundrechte, Sicherheit oder Wohlergehen der Menschen betreffen.

1.2 Aktualisierte Kriterienstruktur

- Welche Kriterien für Trainingsdaten gelten
- Welche Kriterien für Inferenzdaten gelten
- Welche Kriterien für beide Phasen gelten
- Die Betriebskriterien werden ausdrücklich in getrennte Abschnitte Training und Inferenz aufgeteilt
- Übergreifende Governance-Kriterien gelten auf Organisationsebene

1.3 Regulatorische Grundlage

- EU-KI-Verordnung (2024): risikobasierter Rechtsrahmen mit Anforderungen an Hochrisiko-KI, u.

- a. Daten-Governance (Artikel 10), Dokumentation, menschliche Aufsicht und Transparenz
- AIDA Kanadas (Gesetzentwurf C-27): vorgeschlagener Rahmen für Systeme mit hoher Auswirkung, mit Fokus auf Risikobewertung und -minderung, Transparenz und Rechenschaft
- US-Executive Order 14110: umfassender Ansatz für Sicherheit, Datenschutz und Minderung von Verzerrungen (im Januar 2025 aufgehoben, Rahmen bleibt einflussreich)
- NIST-Rahmenwerk für das KI-Risikomanagement: freiwilliger Rahmen mit den Funktionen Steuern, Abbilden, Messen und Verwalten
- ISO/IEC 42001 und 8183: internationale Normen für KI-Managementsysteme und den Datenlebenszyklus

1.4 Unterscheidung Training / Inferenz

Trainingsdaten Daten zum Entwickeln, Trainieren, Validieren und Testen der Modelle: historische Datensätze, gelabelte Daten, synthetische Daten sowie Abstimmungs- und Bewertungsdaten.

Inferenzdaten Daten im operativen Betrieb, wenn das System Vorhersagen, Empfehlungen oder Entscheidungen in der Produktion trifft, einschließlich Echtzeit-Eingaben.

Beide Kriterien, die sowohl für Trainings- als auch für Inferenzdaten gelten, auch wenn die Umsetzung abweichen kann.

2. Bewertungskriterien des Datenlebenszyklus

Die folgenden Kriterien bewerten die Daten-Governance über die Lebenszyklusstufen (ohne Betrieb, siehe Abschnitt 3). Jedes Kriterium gibt an, ob es für Trainingsdaten, Inferenzdaten oder beide gilt.

2.1 Datenerfassung

TRAININGS- UND INFERENZDATEN

Beschaffung von Daten aus verschiedenen Quellen mit den erforderlichen Rechten, Einwilligungen und der Dokumentation.

KRITERIEN

- AC-01 — Rechtsgrundlage der Datenerhebung
- AC-02 — Dokumentation und Herkunft der Quellen
- AC-03 — Verzerrungsbewertung bei der Erhebung
- AC-04 — Schutz von Privatsphäre und Vertraulichkeit
- AC-05 — Einhaltung von Urheber- und Schutzrechten

2.2 Datenaufnahme

TRAININGS- UND INFERENZDATEN

Import und Verarbeitung von Daten mit Qualitätsprüfung, Standardisierung und Sicherheitsmaßnahmen.

KRITERIEN

- IN-01 — Validierung der Datenqualität
- IN-02 — Standardisierung des Datenformats

- IN-03 — Sicherheit bei der Aufnahme
- IN-04 — Qualität von Kennzeichnung und Annotation

2.3 Datenintegration

TRAININGS- UND INFERENZDATEN

KRITERIEN

- IT-01 — Nachverfolgung der Datenherkunft (Lineage)
- IT-02 — Qualitätssicherung der Integration
- IT-03 — Bewertung der Verzerrungsverstärkung
- IT-04 — Stammdatenverwaltung (MDM)

2.4 Datenklassifizierung

TRAININGS- UND INFERENZDATEN

KRITERIEN

- CL-01 — Klassifizierung nach Sensibilität
- CL-02 — Regulatorische Klassifizierung
- CL-03 — Kennzeichnung Training / Inferenz
- CL-04 — Klassifizierung nach Datenqualität

2.5 Datenspeicherung

TRAININGS- UND INFERENZDATEN

KRITERIEN

- ST-01 — Sicherheitskontrollen
- ST-02 — Aufbewahrung und Lebenszyklusverwaltung
- ST-03 — Sicherung und Wiederherstellung
- ST-04 — Audit-Protokollierung und Überwachung
- ST-05 — Resilienz der Speicherinfrastruktur

2.6 Datenpflege

TRAININGS- UND INFERENZDATEN

KRITERIEN

- MA-01 — Verwaltung der Datenaktualität
- MA-02 — Datenbereinigung und -korrektur
- MA-03 — Ausbalancierung und Anreicherung von Datensätzen
- MA-04 — Verwaltung der Schema-Evolution
- MA-05 — Kontinuierliche Datenverbesserung

2.7 Datenweitergabe und -verbreitung

NUR INFERENZDATEN

Gemäß der Zuordnung gelten diese Kriterien nur für Inferenzdaten (operative Weitergabe in der

Produktion).

KRITERIEN

- SH-01 — Governance der Datenweitergabe
- SH-02 — Datenschutz bei der Weitergabe
- SH-03 — Verwaltung externer Auftragsverarbeiter
- SH-04 — Veröffentlichung und Transparenz von Trainingsdaten
- SH-05 — Sicherheit von APIs und Schnittstellen

2.8 Datenarchivierung

TRAININGS- UND INFERENZDATEN

KRITERIEN

- AR-01 — Archivierungsrichtlinie und -verfahren
- AR-02 — Sicherheit und Datenschutz der Archivierung
- AR-03 — Dokumentation der Archivierung
- AR-04 — Langfristige Zugänglichkeit

2.9 Datenentsorgung und -stilllegung

TRAININGS- UND INFERENZDATEN

KRITERIEN

- DR-01 — Sichere Datenlöschung
- DR-02 — Einhaltung des Rechts auf Löschung
- DR-03 — Stilllegung von Modellen
- DR-04 — Dokumentation der Entsorgung
- DR-05 — Migration oder Entsorgung von Altsystemdaten

3. Kriterien des Datenbetriebs

Die Betriebskriterien werden in getrennte Abschnitte für Training und Inferenz aufgeteilt. Jeder Abschnitt enthält dieselben fünf Kriterien (OP-01 bis OP-05), je nach Phase unterschiedlich bewertet.

3.1 Datenbetrieb — Modelltraining

NUR TRAININGSDATEN

Daten-Governance während des aktiven Trainings: Experimente, Hyperparameter-Abstimmung und iterative Entwicklung.

KRITERIEN

- OP-01 — Überwachung der Datenqualität (Training)
- OP-02 — Überwachung der Modellleistung (Training)
- OP-03 — Verfahren der menschlichen Aufsicht (Training)
- OP-04 — Überwachung von Verzerrung und Fairness (Training)
- OP-05 — Echtzeit-Datenvalidierung (Training)

3.2 Datenbetrieb — Modellinferenz

NUR INFERENZDATEN

Daten-Governance im operativen Betrieb, auf Produktionsdaten.

KRITERIEN

- OP-01 — Überwachung der Datenqualität (Inferenz)
- OP-02 — Überwachung der Modelleistung (Inferenz)
- OP-03 — Verfahren der menschlichen Aufsicht (Inferenz)
- OP-04 — Überwachung von Verzerrung und Fairness (Inferenz)
- OP-05 — Echtzeit-Datenvalidierung (Inferenz)

4. Governance- und Ethikkriterien

Diese Kriterien gelten auf Organisationsebene für alle KI-Systeme und Datennutzungen. Sie sind nicht projekt- oder phasenspezifisch und bewerten die Reife und das Engagement der Organisation für verantwortungsvolle KI.

4.1 Organisatorische Governance

- GV-01 — KI-Governance-Struktur
- GV-02 — Risikomanagement-Rahmenwerk
- GV-03 — Compliance-Management

4.2 Dokumentation und Transparenz

- DC-01 — Technische Dokumentation
- DC-02 — Erklärbarkeit und Interpretierbarkeit
- DC-03 — Algorithmische Folgenabschätzungen

4.3 Ethische Grundsätze und Rechenschaft

- ET-01 — Fairness und Nichtdiskriminierung
- ET-02 — Transparenz und Rechenschaft
- ET-03 — Auswirkungen auf die Menschenrechte
- ET-04 — Ökologische Nachhaltigkeit

4.4 Schulung und Sensibilisierung

- TR-01 — Schulung und Kompetenz des Personals
- TR-02 — Organisationskultur

5. Bewertungsmethodik und Bewertung

5.1 Zwei-Rollen-System

- Fachkräfte für Datenethik und -herkunft: sammeln und dokumentieren die Nachweise für jedes Kriterium.

- Bewerter für Datenethik und -herkunft: prüfen die Nachweise und vergeben die Konformitätsbewertungen.

5.2 Bewertung des Konformitätsniveaus (Skala 0 bis 5)

Jedes Kriterium wird auf einer sechsstufigen Skala bewertet:

- 0 Nicht beobachtet** Kein Nachweis der Praxis oder Kontrolle. Kriterium nicht behandelt.
- 1 Initial / ad hoc** Minimale, informelle Umsetzung. Inkonsistente Praxis. Erhebliche Lücken.
- 2 In Entwicklung** Grundlegende Umsetzung mit etwas Dokumentation. Kernelemente vorhanden, Lücken bleiben.
- 3 Definiert / verwaltet** Gut definiert, dokumentiert und konsistent umgesetzt. Erfüllt die grundlegenden regulatorischen Anforderungen.
- 4 Gemessen / optimiert** Umfassend, mit aktiver Überwachung und kontinuierlicher Verbesserung. Übertrifft die Grundanforderungen.
- 5 Führende Praxis** Vorbildliche, branchenführende Umsetzung. Innovative Ansätze mit nachgewiesener Wirksamkeit.

5.3 Gewichtung nach Kritikalität

Jedem Kriterium wird eine Kritikalitätsstufe zugewiesen:

- Kritisch (C): Gewichtung 40 % — eine Bewertung < 3 blockiert die Bereitstellung
- Hoch (H): Gewichtung 30 % — eine Bewertung < 2 erfordert einen Maßnahmenplan
- Mittel (M): Gewichtung 20 % — eine Bewertung < 2 löst einen Verbesserungsfahrplan aus
- Niedrig (L): Gewichtung 10 % — Fokus auf kontinuierliche Verbesserung

5.4 Gesamtakkreditierungsstufen

- Stufe 5 (Vorbildlich): Gesamtbewertung $\geq 4,5$, kein kritisches Kriterium < 4, organisatorische Exzellenz
- Stufe 4 (Fortgeschritten): Gesamtbewertung $\geq 4,0$, kein kritisches Kriterium < 3, starke Leistung
- Stufe 3 (Kompetent): Gesamtbewertung $\geq 3,0$, kein kritisches Kriterium < 3, erfüllt alle regulatorischen Anforderungen
- Stufe 2 (In Entwicklung): Gesamtbewertung $\geq 2,0$, Maßnahmenplan erforderlich
- Stufe 1 (Nicht konform): Gesamtbewertung < 2,0, sofortige Maßnahmen, Bereitstellung kann blockiert werden

6. Hinweise zur Umsetzung

6.1 Arbeitsablauf der Fachkraft

- Abgrenzung: KI-System und anwendbare Kriterien bestimmen (Training/Inferenz)
- Planung: einen Zeitplan für die Nachweiserhebung erstellen
- Nachweiserhebung: Dokumentation sammeln, Beteiligte befragen, Systeme testen
- Dokumentation: die Nachweisvorlagen für jedes Kriterium ausfüllen
- Lückenermittlung: Bereiche mit fehlenden Nachweisen benennen

- Qualitätsprüfung: Selbstprüfung auf Vollständigkeit vor der Einreichung
- Einreichung: das Nachweispaket an die Bewerter übergeben

6.2 Arbeitsablauf des Bewerter

- Nachweisprüfung: Vollständigkeit und Qualität der Nachweise prüfen
- Verifizierung: durch Stichproben, Tests oder Interviews validieren
- Bewertung: Bewertungen von 0 bis 5 anhand von Nachweisen und Leitfaden vergeben
- Dokumentation: Feststellungen, Begründungen und Empfehlungen festhalten
- Kalibrierung: Peer-Review der Bewertungen zur Konsistenz
- Bericht: einen Bericht mit Bewertungen und Feststellungen erstellen
- Nachverfolgung: die Behebung der festgestellten Lücken verfolgen

6.3 Auslöser für eine Neubewertung

- Jährliche Neubewertung aller bereitgestellten Hochrisikosysteme
- Wesentliche Systemänderungen (neue Datenquellen, Algorithmusänderungen, Erweiterung der Anwendungsfälle)
- Regulatorische Änderungen, die das KI-System betreffen
- Wesentliche Vorfälle oder entdeckte Verzerrungen
- Übergang vom Pilot- zum Produktionsbetrieb

Anhang A: Kriterienübersicht nach Phase

Schnellreferenz, welche Kriterien für welche Phase gelten:

Kriterien für Trainingsdaten

- Alle Kriterien der Stufen 1 bis 6 (von der Erfassung bis zur Speicherung)
- Datenbetrieb — Training (OP-01 bis OP-05)
- Datenpflege (MA-01 bis MA-05)
- Datenarchivierung (AR-01 bis AR-04)
- Datenentsorgung und -stilllegung (DR-01 bis DR-05)

Kriterien für Inferenzdaten

- Alle Kriterien der Stufen 1 bis 6 (von der Erfassung bis zur Speicherung)
- Datenbetrieb — Inferenz (OP-01 bis OP-05)
- Datenpflege (MA-01 bis MA-05)
- Datenweitergabe und -verbreitung (SH-01 bis SH-05)
- Datenarchivierung (AR-01 bis AR-04)
- Datenentsorgung und -stilllegung (DR-01 bis DR-05)

Organisatorische (übergreifende) Kriterien

- Organisatorische Governance (GV-01 bis GV-03)

- Dokumentation und Transparenz (DC-01 bis DC-03)
- Ethische Grundsätze und Rechenschaft (ET-01 bis ET-04)
- Schulung und Sensibilisierung (TR-01 bis TR-02)

Anhang B: Glossar

Trainingsdaten

Daten zum Entwickeln, Trainieren, Validieren und Testen der Modelle im Entwicklungszyklus.

Inferenzdaten

Daten im operativen Betrieb, wenn das System Vorhersagen, Empfehlungen oder Entscheidungen trifft.

Datenlebenszyklus

Der vollständige Weg der Daten von der Erfassung bis zur Entsorgung über 10 Stufen.

Hochrisiko-KI-System

KI-System mit erheblichen Risiken für Gesundheit, Sicherheit oder Grundrechte (gemäß EU-KI-Verordnung).

Verzerrung (Bias)

Systematische Fehler oder Ungerechtigkeiten in KI-Ausgaben, die zu diskriminierender Behandlung nach geschützten Merkmalen führen.

Fairness-Metriken

Quantitative Maße für Fairness: demografische Parität, ausgeglichene Chancen oder prädiktive Parität.

Datendrift

Veränderung der statistischen Eigenschaften der Eingabedaten im Zeitverlauf, die das Modell verschlechtern kann.

Konzeptdrift

Veränderung der Beziehung zwischen Eingaben und Ausgaben im Zeitverlauf, die ein Neutraining erfordert.

Modellkarte

Strukturierte Dokumentation der Eigenschaften, des Verwendungszwecks, der Grenzen und der Leistung eines Modells.

Datensatzkarte

Strukturierte Dokumentation der Eigenschaften, Erhebungsmethoden und bekannten Grenzen eines Datensatzes.

Differenzielle Privatsphäre

Mathematischer Rahmen zur Begrenzung des Privatsphärenverlusts bei der Analyse von Datensätzen.

Erklärbarkeit (XAI)

Die Fähigkeit, verständliche Erklärungen für Entscheidungen und Verhalten eines KI-Systems zu liefern.

Anhang C: Referenzen

- Europäische Union. (2024). Verordnung (EU) 2024/1689 über künstliche Intelligenz (KI-Verordnung).
- Regierung von Kanada. (2022). Gesetzentwurf C-27, Digital Charter Implementation Act, 2022 (AIDA).
- Biden-Regierung. (2023). Executive Order 14110 über sichere und vertrauenswürdige Entwicklung und Nutzung von KI.
- NIST. (2023). Rahmenwerk für das KI-Risikomanagement (AI RMF 1.0). NIST AI 100-1.
- NIST. (2024). Profil für generative KI. NIST AI 600-1.
- ISO/IEC. (2023). ISO/IEC 42001:2023 — Künstliche Intelligenz — Managementsystem.
- ISO/IEC. (2024). ISO/IEC 8183:2024 — Rahmen für den Datenlebenszyklus.
- Europäische Kommission. (2016). Datenschutz-Grundverordnung (DSGVO) — (EU) 2016/679.
- Datenschutzbeauftragter von Kanada. PIPEDA.
- Bundesstaat Kalifornien. (2018). California Consumer Privacy Act (CCPA) und CPRA.

Versionsverlauf

- Version 1.0: erstes umfassendes Rahmenwerk
- Version 2.0: Aktualisierung nach der offiziellen Kriterienzuordnung, mit expliziten Unterscheidungen Training/Inferenz/Organisatorisch