

DIGITALE ETHIK: ORDNUNG IM NORMEN-DSCHUNGEL

1 Einleitung

In den letzten Jahren entstanden zahlreiche freiwillige Selbstregulierungs-Initiativen in verschiedenen Regionen, Ländern und Branchen. Das Spektrum reicht vom Engagement der einzelnen Unternehmen und Verbände über die Gründung gemeinsamer Organisationen (z.B. Partnership for AI) bis hin zu Selbstverpflichtungen gegenüber dem Staat (z.B. Verhaltenskodex der EU gegen Desinformation). Ziel ist es, vertrauenswürdige digitale Produkte und Dienstleistungen zu entwickeln und dies transparent zu machen. So lässt sich das Vertrauen der Bevölkerung in datenbasierte Innovationen stärken. Dabei kristallisieren sich grob zwei Ansätze heraus:

- **Orientierung an Prinzipien:** Unternehmen und Organisationen befolgen Kodizes und Ethik-Richtlinien, in denen bestimmte Anforderungen oder Grundsätze festgehalten sind. Dabei geht es zum Beispiel um Schadensvermeidung, Autonomie, Gerechtigkeit, Fairness, Privatsphäre und Transparenz. Die Orientierung an solchen Prinzipien erfreut sich grosser Beliebtheit. Sie werden von einzelnen Unternehmen und Organisationen für den eigenen Gebrauch erarbeitet, basierend auf ihren Werten und Ambitionen. Eine Alternative sind Ethik-Richtlinien von Netzwerken oder Verbänden, die sich auf die spezifische Situation eines Unternehmens adaptieren lassen (z.B. Swico Ethik Charta). Denn bei der Orientierung an Prinzipien ist wichtig, dass diese zu den Werten und zur Kultur des Unternehmens passen. Erst dann entfalten sie ihre Wirkung.
- **Labels, Standards und Normen für Produkte:** Für datenbasierte Produkte und Dienstleistungen werden eine Reihe von messbaren Kriterien festgelegt, um die Ethik-Richtlinien zu operationalisieren. Diese Kriterien bilden gleichzeitig die Voraussetzung für Zertifizierungen und Audits. Die Lösungsvorschläge für mögliche Labels, Standards und Normen sind vielfältig und aufgrund der hohen Komplexität befindet sich die Mehrzahl noch in der Entwicklungsphase.

Für Unternehmen der ICT-Branche in der Schweiz lohnt sich die Orientierung an Prinzipien, zum Beispiel basierend auf der Swico Ethik Charta. Diese ist pragmatisch aufgebaut und hilft anhand von insgesamt 24 Bekenntnissen, datenbasierte Produkte und Dienstleistungen menschenzentriert und fair auszugestalten.

Gleichzeitig sollte die Entwicklung neuer Labels, Standards und Normen für datenbasierte Angebote im Auge behalten werden. Sie schaffen Klarheit und lassen sich in der Kommunikation gegenüber Kundinnen und Kunden sowie weiteren Stakeholdern nutzen. Gerade für grosse Unternehmen ist die kommunikative Positionierung als vertrauenswürdiges Unternehmen attraktiv, auch angesichts neuer Regulierungen. Unternehmen der ICT-Branche können sich freiwillig für gewisse Labels, Standards und Normen entscheiden. Denkbar ist auch, dass sie von diesen als Lieferant betroffen sind im Rahmen neuer Kundenaufträge.

2 Regulatorische Einbettung

Ein wichtiger Treiber für Labels, Standards und Normen sind, neben dem Bedürfnis nach Orientierung, aktuelle sowie geplante Regulierungen. Dazu gehören insbesondere:

- **Datenschutz:** Das [verschärfte Datenschutzgesetz](#) in der Schweiz, welches voraussichtlich am 1. September 2023 in Kraft tritt, wird die Sensibilität für den verantwortungsvollen Umgang mit Daten in der Schweiz weiter erhöhen. Ausserdem sind viele Unternehmen von der Datenschutz-Grundverordnung (DSGVO) für den europäischen Binnenmarkt betroffen.
- **Künstliche Intelligenz (KI):** Zurzeit ist kein Gesetz zur Regulierung von KI in der Schweiz geplant, doch viele Unternehmen werden – analog zur DSGVO – vom neuen [Artificial Intelligence Act \(AI Act\)](#) der EU betroffen sein. KI-Systeme, die in der EU verwendet werden, sollen sicher, transparent, ethisch, unparteiisch und unter menschlicher Kontrolle sein. Sie werden daher nach ihrem Risiko für Gesundheit, Sicherheit und Grundrechte eingeteilt. Geplant sind vier Risikostufen: KI-Systeme mit unzulässigem Risiko (Verbot), mit hohem Risiko, mit begrenztem Risiko sowie mit minimalem Risiko. Der AI Act befindet sich derzeit im parlamentarischen Prozess. Er wird frühestens in der zweiten Hälfte 2024 in Kraft treten können.

Angesichts dieser Regulierungen steigt die Attraktivität für Unternehmen, sich an gewisse Labels und Standards (vgl. Kapitel 3) auf dem europäischen Markt zu halten. Gleiches gilt für Normen (vgl. Kapitel 4). Sie wirken ergänzend, indem sie eher technische Anforderungen an Produkte und Verfahren festlegen.

3 Ausgewählte Labels und Standards

Nr.	Standards und Normen	Ziel, Inhalt, Trägerschaft	Adressat	Eignung für Swico-Mitglieder
1	Privacy Icons	Unternehmen schaffen Orientierung bei der Datenschutzerklärung, indem zentrale Informationen zusätzlich visualisiert werden (Piktogramme) Trägerschaft: Verein Privacy Icons	Unternehmen in der Schweiz	Ja: Geringer Aufwand
2	Data Fairness Label	Unternehmen machen ihren verantwortungsvollen Umgang mit grossen Datenmengen sichtbar (Selbstverpflichtung + Label) Trägerschaft: Verband Swiss Insights	Unternehmen in der Schweiz und im Ausland, idealerweise mit einer Datenabteilung von mind. einer Person (1 FTE)	Ja: Pragmatischer Ansatz mit Wirkung
3	Digital Trust Label	Bei Kunden, die digitale Anwendungen nutzen, wird Vertrauen und Transparenz geschaffen, indem einzelne Anwendungen zertifiziert sind (unabhängiges Audit + Label) Trägerschaft: Stiftung Swiss Digital Initiative	Unternehmen weltweit, bestimmte digitale Anwendungen	Eher nein: Aufwändig und teuer
4	AI Trust Label (in Arbeit)	Die Vertrauenswürdigkeit eines KI-Produkts wird beschrieben und gemessen, was Produktvergleiche erlaubt (Selbstverpflichtung oder Audit + Label) Trägerschaft: Konsortium mit Verband der Elektrotechnik Elektronik Informationstechnik (VDE) und weiteren Organisationen	Unternehmen weltweit	Tendenz Ja
5	KI Gütesiegel	Unternehmen machen sichtbar, dass sie	Unternehmen in Deutschland	Nein: Fehlende Glaubwürdigkeit

Nr.	Standards und Normen	Ziel, Inhalt, Trägerschaft	Adressat	Eignung für Swico-Mitglieder
		ethisch vertretbare Services und Produkte entwickeln bzw. einsetzen (Selbstverpflichtung + Label) Trägerschaft: KI Bundesverband		
6	Kriterienkatalog SustAIn	Sensibilisierung von Wirtschaft, Gesellschaft und Politik, KI-Systeme einer Nachhaltigkeitsbewertung zu unterziehen mittels Kriterienkatalog (Hilfestellung) Trägerschaft: Institut für ökologische Wirtschaftsforschung (iöw) in Kooperation mit der Technischen Universität Berlin und AlgorithmWatch	Unternehmen weltweit	Eher ja: Sehr umfangreich
7	Standard IEEE 7000	Unternehmen entwickeln KI-Produkte so, dass ethische Überlegungen berücksichtigt und Risiken reduziert werden (Standard / Hilfestellung) Trägerschaft: Institute of Electrical and Electronics Engineers (IEEE)	Unternehmen weltweit	Ja: Zielgruppe sind auch kleine Unternehmen
8	KI-Prüfkatalog Fraunhofer	Leitfaden für Unternehmen und Prüfer, um die Vertrauenswürdigkeit eines KI-Systems zu beurteilen (Hilfestellung) Trägerschaft: Fraunhofer-Institut	Unternehmen weltweit	Eher ja: Sehr umfangreich

Hinweis: Die Beurteilung der Eignung für Swico-Mitglieder basiert auf generellen Annahmen. Davon kann die Eignung für ein spezifisches ICT-Unternehmen abweichen, je nach Ambitionsniveau und Geschäftsmodell.

3.1 Privacy Icons

Um einzelne Aspekte der Datenschutzerklärung einfach verständlich abzubilden, gibt es die «Privacy Icons». Ziel dieser Piktogramme ist es, dass sich Kunden schnell einen Überblick zu zentralen Aspekten der Datenbearbeitung verschaffen können. Sie ergänzen die Datenschutzerklärung auf der Website. Die grafischen Elemente, es stehen 19 Privacy Icons zur Verfügung, stehen allen Schweizer Unternehmen zur Verfügung. Sie decken das Schweizer Datenschutzgesetz ab. Herausgeberin ist der Verein Privacy Icons, ein Zusammenschluss von Schweizer Unternehmen mit dem Ziel, Datenschutzhinweise verständlicher zu machen und die Transparenz im Datenschutz zu erhöhen.

Würdigung: Die «Privacy Icons» eignen sich für alle Unternehmen. Sie können kostenlos genutzt werden und der Aufwand für die Implementierung dieser Piktogramme auf der Website ist überschaubar. Fraglich ist, ob die beabsichtigte Vereinfachung erreicht wird.

3.2 Data Fairness Label

Mit dem Label «Data Fairness» verpflichten sich Unternehmen (oder ihre Datenabteilung) zu einem transparenten und verantwortungsvollen Umgang mit Daten. Das Label ist ein sichtbares Bekenntnis zu den Werten Transparenz und Fairness nach innen und aussen. Voraussetzung ist, dass die wichtigsten Merkmale der verwendeten Datensätze (Data Sheets) oder die Art der verwendeten Datenmodelle inklusive Verwendungszweck (Model Cards) über ein Online-Tool erfasst werden. Auch die Prozesse werden online dokumentiert. Das Label wird vom Verband Swiss Insights herausgegeben und steht inländischen und ausländischen Unternehmen zur Verfügung. Eine Zertifizierung ist in Zukunft denkbar, sofern dies einem Bedürfnis der Unternehmen entspricht.

Würdigung: Das Label wird von einem glaubwürdigen Verband herausgegeben und eignet sich für alle Unternehmen, die mit grösseren Datensätzen arbeiten. Das Label wurde von erfahrenen Praktikern entwickelt. Wer das Label nutzt, kann vom Erfahrungsaustausch in der Label-Community profitieren und die Weiterentwicklung des Labels aktiv mitgestalten.

3.3 Digital Trust Label

Mit dem «Digital Trust Label» kennzeichnet die Swiss Digital Initiative, ein Spin-off von Digital Switzerland, die Vertrauenswürdigkeit digitaler Anwendungen. Das Spektrum reicht von einfachen Websites bis hin zu Internet-Infrastrukturdiensten und Online-Plattformen. Ziel ist, dass sich Nutzer sicher fühlen, wenn sie digitale Anwendungen nutzen. Das Label eignet sich vor allem für digitale Anwendungen, bei denen eine Vielzahl an sensitiven Informationen (z.B. Nutzerdaten) verwendet werden und/oder ein automatisierter Entscheidungsprozess stattfindet. Zu den besonders betroffenen Industrien gehören der öffentliche Sektor, Gesundheitswesen, Medien, Personalwesen, Bildung sowie Banken und Versicherungen. Unternehmen aus dem In- und Ausland können ihre digitalen Produkte mit dem «Digital Trust Label» zertifizieren lassen, wenn sie 35 Kriterien in den vier Bereichen Sicherheit, Datenschutz, Verlässlichkeit und User Interaktion erfüllen, was von einem Audit bestätigt werden muss. Der Kriterienkatalog baut auf bestehenden Normen wie ISO 27001, ISO 22301 und der Datenschutz-Grundverordnung (DSGVO) auf. Ziel ist ein internationales Label, das Unternehmen eine Differenzierung am Markt ermöglicht.

Würdigung: Das Label eignet sich vor allem für Grossunternehmen, die sich gegenüber ihrer Kundschaft als vertrauenswürdig positionieren möchten. Der Prozess ist sehr aufwändig und das Label teuer. ICT-Unternehmen könnten Unternehmen, die das Label erwerben möchten, bei der Umsetzung unterstützen.

3.4 AI Trust Label (in Arbeit)

Mit dem «AI Trust Label» lässt sich die Vertrauenswürdigkeit und die Einhaltung bestimmter Werte von KI-Produkten belegen. Dies geschieht über die Operationalisierung der Werte Transparenz, Verantwortlichkeit, Schutz der Privatsphäre, Fairness und Verlässlichkeit. Um die zur Verfügung stehenden Produkte vergleichen zu können, orientiert sich das Label am Energieeffizienzlabel für Elektrogeräte im Sinne einer abgestuften Kennzeichnung relevanter Eigenschaften eines KI-Systems. Das Label soll, je nach Anwendungsfall, eine Selbsterklärung beziehungsweise eine Zertifizierung durch unabhängige Instanzen erlauben. Das «AI Trust Label» basiert auf einer Spezifikation des VDE (Verband der Elektrotechnik Elektronik Informationstechnik) und wird von einem Konsortium aus Wirtschaft, Wissenschaft und Zivilgesellschaft unterstützt. Ziel ist ein international anerkannter Standard, der mit dem AI Act der EU kompatibel ist und eine Differenzierung am Markt ermöglicht.

Würdigung: Das «AI Trust Label» basiert auf konzeptionellen Grundlagen, die von der AI Ethics Impact Group entwickelt wurden, initiiert von der Bertelsmann Stiftung und dem Verband der Elektrotechnik Elektronik Informationstechnik (VDE). Das Label ist vielversprechend aber noch nicht auf dem Markt. Die Präsentation weiterer Details zum Label ist Mitte 2022 geplant.

3.5 KI Gütesiegel

Das «KI Gütesiegel» ist ein Indikator dafür, dass die Produkte und Dienstleistungen eines Unternehmens branchenübliche Qualitätsanforderungen erfüllen und nach anerkannten Regeln der Technik entwickelt werden. Mit dem Gütesiegel soll die Entwicklung von ethisch vertretbaren Services und Produkten gefördert werden, basierend auf den vier Gütekriterien Ethik, Unvoreingenommenheit (Bias), Transparenz, Sicherheit und Datenschutz. Deutsche Unternehmen, die sich schriftlich verpflichten, diese Werte und die dazu passenden Hinweise bzw. Empfehlungen zu berücksichtigen, können das «KI Gütesiegel» in der Kommunikation nutzen. Herausgeberin ist der KI Bundesverband.

Würdigung: Das Label wurde bereits 2019 lanciert, seither aber nicht weiterentwickelt. Die Werte und Kriterien sind vage und unvollständig, was den Eindruck erweckt, dass es eher ein Angebot für Mitglieder des KI Bundesverbandes ist, um sich positionieren zu können. Ob es Unternehmen in Deutschland gibt, die das «KI Gütesiegel» einsetzen, ist nicht klar. In vielen Studien fehlt dieses Gütesiegel, wahrscheinlich aufgrund der fehlenden Glaubwürdigkeit und Relevanz.

3.6 Kriterienkatalog SustAIIn

Ziel des Forschungsprojekts «SustAIIn – Nachhaltigkeitsindex für Künstliche Intelligenz» ist es, den gesellschaftlichen Diskurs zur Nachhaltigkeit von KI-Systemen zu stärken sowie eine systematische Nachhaltigkeitsbewertung zu ermöglichen. Dafür werden die sozialen, ökologischen und ökonomischen Auswirkungen entlang des gesamten KI-Lebenszyklus sowie die organisationale Einbettung dieser Systeme betrachtet. Das entwickelte Kriterien- und Indikatorenset besteht aus dreizehn Nachhaltigkeitskriterien (z.B. Energieverbrauch, Kulturelle Sensibilität, Arbeitsbedingungen & Arbeitsplätze) mit entsprechenden Indikatoren sowie fünf Querschnitts-Indikatoren (z.B. Code of Conduct, Risikomanagement, Stakeholderanalyse & Beteiligung). Erarbeitet wurde die Nachhaltigkeitsbewertung von KI-Systemen vom Institut für ökologische Wirtschaftsforschung (iöw) in Kooperation mit der Technischen Universität Berlin und AlgorithmWatch.

Würdigung: Das Forschungsprojekt schliesst eine Lücke, indem die Nachhaltigkeit von KI-Systemen als Bestandteil einer vertrauenswürdigen KI thematisiert und messbar gemacht wird. Das Kriterien- und Indikatorenset ist sehr umfangreich. Für ICT-Unternehmen könnten gewisse Kriterien wertvoll sein bei Kunden, die an diesem Mehrwert im Rahmen ihrer Nachhaltigkeitsstrategie interessiert sind.

3.7 Standard IEEE 7000

Der Standard IEEE 7000 ist international ausgerichtet. Er richtet sich explizit an alle Nationen, Kulturen und Branchen. Dabei handelt es sich um ein sehr umfangreiches Framework mit dem Ziel, ethische KI-Systeme zu entwickeln und zu betreiben, bei denen der Mensch im Mittelpunkt steht. Der Standard verfolgt einen holistischen Ansatz, bei dem die Werte der Nutzer während dem Designprozess berücksichtigt werden. Dabei geht es um den gesamten Lebenszyklus eines Produkts, wobei die Entwicklungsphase besonders stark gewichtet wird. Verantwortlich für den Standard ist das Institute of Electrical and Electronics Engineers (IEEE), ein renommierter internationaler Berufsverband. Der Verband ist überzeugt, dass verantwortungsvolle Innovationen im Zeitalter der Algorithmen eine werteorientierte Methodik erfordern, die das traditionelle Systems Engineering ergänzt. Der Standard IEEE 7000 kann Unternehmen jeder Grösse helfen, Produkte mit einem differenzierten Nutzenversprechen und mit weniger Risiko zu entwickeln.

Würdigung: Der Standard ist international ausgerichtet und grundsätzlich für alle Unternehmen geeignet, insbesondere in der Entwicklungsphase von KI-Anwendungen. Aufgrund seines Umfangs dürfte ein gewisses Fachwissen von Vorteil sein, um diese Hilfestellung mit vernünftigen Aufwand nutzen zu können.

3.8 KI-Prüfkatalog Fraunhofer

Der Leitfaden zur Gestaltung vertrauenswürdiger Künstlicher Intelligenz ist für Entwickler eine wertvolle Richtschnur, um neue KI-Anwendungen systematisch vertrauenswürdig zu gestalten. Ausserdem kann der Leitfaden als Prüfkatalog genutzt werden, um KI-Anwendungen strukturiert auf Vertrauenswürdigkeit zu untersuchen. Die Arbeit erfolgt in beiden Fällen mit Blick auf sechs Dimensionen von Vertrauenswürdigkeit: Fairness, Autonomie und Kontrolle, Transparenz, Verlässlichkeit, Sicherheit sowie Datenschutz. Der Leitfaden ist sehr umfassend und detailliert. Herausgeberin ist das renommierte Fraunhofer-Institut.

Würdigung: Der Leitfaden deckt sechs Dimensionen von Vertrauenswürdigkeit detailliert ab und ist vor allem für Prüfer geeignet. Für ICT-Unternehmen kann er eine wertvolle Hilfestellung sein, etwa wenn (grössere) Kunden spezifische Bedürfnisse realisiert haben möchten.

4 Die Rolle technischer Anforderungen

Der AI Act der EU sieht vor, dass Anbieter von KI-Systemen mit hohem Risiko künftig ein Verfahren durchlaufen, das die Konformität ihrer Produkte mit den gesetzlichen Vorgaben überprüft (u.a. an Transparenz, Robustheit und Genauigkeit). Dabei spielen harmonisierte europäische Normen eine zentrale Rolle. Etablierte Institutionen aus dem Bereich Normierung, Prüfung und Sicherheitskontrolle aus

Deutschland sind bereits aktiv. Ihre Lösungsvorschläge tragen zum Ziel der Bundesregierung bei, eine vertrauenswürdige «KI made in Germany» zu entwickeln und als weltweiten Massstab zu etablieren.

Interessant für ICT-Unternehmen in der Schweiz sind folgende laufende Arbeiten im Hinblick auf technische Anforderungen an KI-Systeme:

- **ISO-Norm:** Die internationale Organisation für Normierung hat den neuen Standard ISO/IEC TR 24028 zur Vertrauenswürdigkeit von KI-Systemen entwickelt. Ziel ist die Schaffung von Vertrauen durch Transparenz, Kontrolle und Erklärbarkeit sowie der Gewährleistung der Verfügbarkeit, Resilienz, Zuverlässigkeit, Richtigkeit, Sicherheit und des Datenschutzes durch spezifische risikominimierende Ansätze.
- **DIN-Norm:** Das Deutsche Institut für Normung (DIN) hat mehrere Dokumente im Kontext von Künstlicher Intelligenz veröffentlicht. Ein Ziel ist zum Beispiel, dass die Qualität von KI-Anwendungen entlang des Lebenszyklus sichergestellt wird, basierend auf den Qualitätsmerkmalen Funktionalität, Robustheit und Verständlichkeit (DIN SPEC 92001-1). Vertiefte Erläuterungen zur Risikobewertung von KI sind Gegenstand der zweiten Spezifikation (92001-2). Weitere Spezifikationen sind in Arbeit. Ausserdem hat DIN im Auftrag der deutschen Bundesregierung das Projekt «KI-Tauglichkeit von Normen» lanciert. Bei mehr als 30'000 Normen wird überprüft, ob sie für den Einsatz von Künstlicher Intelligenz vorbereitet sind und wo Überarbeitungen nötig sind. Die Ergebnisse sollen Ende 2023 vorliegen.
- **KI-Prüfkatalog TÜV:** Der TÜV Deutschland arbeitet an einem «KI-TÜV». Ähnlich wie bei Motorfahrzeugen sollen KI-Anwendungen regelmässigen Kontrollen unterzogen werden. Abgedeckt werden im KI-Lebenszyklus die Themen Datenschutz und Datensicherheit, Robustheit und Sicherheit, Transparenz und Rechenschaftspflicht. Die Prüfung einer KI-Anwendung soll durch eine unabhängige Partei erfolgen.
- **Normungsroadmap KI:** Ziel ist die Entwicklung eines Handlungsrahmens für die Normung und Standardisierung von KI-Anwendungen. Die Erarbeitung der Roadmap erfolgt im Auftrag der deutschen Bundesregierung. Federführend sind die beiden Normungsorganisationen DIN und DKE. Ein erster Bericht wurde 2020 publiziert. Der zweite Bericht soll diese Ergebnisse fortschreiben und weiterentwickeln. Dieser neue Bericht soll Ende 2022 vorliegen.

5 Ausblick

Das Umfeld und die Anzahl an Labels, Standards und Normen bleiben äusserst dynamisch, weshalb es sich lohnt, deren Relevanz für die ICT-Branche in der Schweiz weiterhin zu beobachten. Hinzu kommt, dass gewisse technische Anforderungen an KI-Systeme erst auf den Markt kommen, wenn der AI Act der EU verabschiedet bzw. in Kraft gesetzt wird.

Wer sich vertieft mit dem Thema auseinandersetzen möchte: In zwei Studien sowie einem Inventar für Ethik-Richtlinien werden zahlreiche Richtlinien, Labels, Standards und Normen aufgeführt. Einige Informationen sind allerdings bereits veraltet und Neues fehlt zum Teil. Dennoch kann sich der Blick in diese Quellen lohnen:

- AlgorithmWatch (Stand April 2020): [AI Ethics Guidelines Global Inventory](#)
- RAND Europe (2022): [Labelling initiatives, codes of conduct and other self-regulatory mechanisms for artificial intelligence applications. From principles to practice and considerations for the future](#)
- Swiss Digital Initiative (2021): [Labels and Certifications for the Digital World. Mapping the International Landscape](#)

Stand: 17. August 2022

Für Rückfragen:

Judith Bellaiche

SWICO