

# ETHIK-KODEX

für datenbasierte Wertschöpfung



**EMPFEHLUNGEN**

# ÜBERSICHT

Der «Ethik-Kodex für datenbasierte Wertschöpfung» besteht aus den folgenden Dokumenten: 1) Übersicht; 2) Grundlagen; 3) Empfehlungen; 4) Implementierung; 5) Kontext.

Im Dokument **EMPFEHLUNGEN** werden konkrete Ratschläge für eine ethische datenbasierte Wertschöpfung gegeben, strukturiert mittels der vier Phasen des Daten-Lebenszyklus. Diese werden sowohl anhand der drei ethischen Grundorientierungen als auch der drei prozeduralen Werte gruppiert, so dass pro Schritt sechs Gruppen von Empfehlungen resultieren. Dazu wird jeweils pro Gruppe erläutert, welche Absichten mit den Empfehlungen verfolgt werden, bevor dann die konkreten Empfehlungen folgen. Zu jeder Gruppe wird anhand eines Beispiels erläutert, wie eine konkrete Umsetzung der Empfehlung aussehen kann. Erläuterungen zum Daten-Lebenszyklus und zu den Grundorientierungen finden sich im Dokument «Grundlagen». Die Empfehlungen liegt in den Sprachen Deutsch, Englisch, Französisch und Italienisch vor.

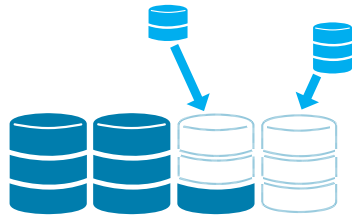
## IMPRESSUM

Der «Ethische Kodex für datenbasierte Wertschöpfung» wurde von der Expertengruppe «Datenethik» der Swiss Alliance for Data-Intensive Services erstellt. Redaktionsteam: Markus Christen, Christoph Heitz, Tom Kleiber, Michele Loi (Projektleitung). Französische Übersetzung: Jean-Gabriel Piguet. Grafik: Ana Nicolasa Caduff, Hochschule Luzern CC Management & Law. Stand Jahr 2020.

© Swiss Alliance for Data-Intensive Services, 2020.

ISBN 978-3-9522703-0-1; [www.data-service-alliance.ch/codex](http://www.data-service-alliance.ch/codex)

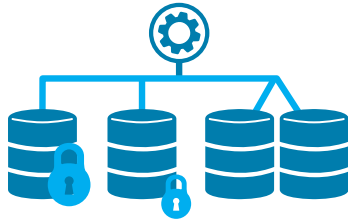
Lizenz: Namensnennung 4.0 International (CC BY 4.0).



## **SCHRITT 1** **Datenerzeugung und -akquirierung**

Ethische Grundorientierungen  
Prozedurale Werte

Seite 4



## **SCHRITT 2** **Datenspeicherung und -management**

Ethische Grundorientierungen  
Prozedurale Werte

Seite 10



## **SCHRITT 3** **Datenanalyse und Wissensgenerierung**

Ethische Grundorientierungen  
Prozedurale Werte

Seite 16



## **SCHRITT 4** **Produkte und Dienstleistungen**

Ethische Grundorientierungen  
Prozedurale Werte

Seite 22

# SCHRITT 1: DATENERZEUGUNG UND -AKQUIRIERUNG

## ETHISCHE GRUNDORIENTIERUNGEN

### SCHADENVERMEIDUNG

Die Empfehlungen zur Grundorientierung Schadenvermeidung sollen sicherstellen, dass durch die Art der Datenerhebung dem Nutzer kein Schaden entsteht. Fragen zu einem möglichen Schaden können sich dann stellen, wenn die Interaktion mit dem Nutzer mittels Erkenntnissen aus der Verhaltensforschung so ausgestaltet werden, dass diese dazu bewegen werden, mehr Daten von sich preiszugeben; man spricht dann von so genannten «Nudges». Ist ein Benutzerschnittstellen-Design willentlich darauf ausgelegt, einen User zur Preisgabe von Informationen zu bringen, die dessen Interessen entgegenlaufen, spricht man von Dark Pattern.



#### Empfehlungen

- Vermeiden Sie es, die Trägheit, mangelnde Aufmerksamkeit oder andere psychologische Schwächen von Personen dahingehend auszunutzen, mehr persönliche Daten von diesen zu erhalten als notwendig sind, um Ihre Dienstleistung zu erbringen.
- Überprüfen Sie Webdesign und Benutzererfahrung auf Dark Patterns und entfernen Sie diese.
- Wenn Sie Nudges verwenden, dann sollten Sie im Voraus nachweisen oder nachträglich dokumentieren, dass diese auch den Interessen des Nutzers dienen (z.B. Erhöhung der Cybersicherheit, Verbesserung des bereitgestellten Dienstes).

### BEISPIEL - DARK PATTERN VERMEIDEN

Dark Patterns sollen den Nutzer irreführen und – im schlechtesten Fall – zu einer Handlung verleiten, die er so gar nicht durchführen wollte. Um dieses «Fehlverhalten» auszulösen, werden erlernte Verhaltensmuster ausgenutzt und User durch geschickt platzierte Buttons oder irreführende Drop-Down Menüs manipuliert. Ein Beispiel ist ein Unternehmen, das nach einer Reise-Buchung gleich auch eine entsprechende Versicherung dem Kunden verkaufen wollte. Die Option «keine Versicherung» findet sich gut versteckt unter «K» in einer alphabetischen Liste von Ländern, in denen die Versicherung jeweils gelten soll. Stattdessen sollte man bei Interfaces die Usability durch Logik unterstützen. Das bedeutet, dass man im konkreten Fall die Grundentscheidung «Ja» versus «Nein» von den möglichen Optionen, die ein «Ja» beinhaltet, abtrennt.



## GERECHTIGKEIT

Die Empfehlungen zur Grundorientierung Gerechtigkeit sollen ein gewisses Mass an Gegenseitigkeit sicherstellen, so dass beim Sammeln von Daten den Betroffenen Vorteile geboten werden, die in einem angemessenen Verhältnis zu den von Ihnen angeforderten Daten stehen.

### Empfehlungen

- Verdeutlichen Sie Ihren Kunden den Gegenwert, den Sie durch die Bekanntgabe ihrer Daten erhalten sowohl für sich selbst als auch gegebenenfalls für eine grössere Gemeinschaft.
- Wenn die Erhebung bestimmter Daten aus gesetzlichen Anforderungen herrührt, machen Sie dies deutlich.



### BEISPIEL - ELEKTRONISCHES PATIENTENDOSSIER (EPD)

Eine Ärztin mit eigener Praxis hat sich beim EPD-System angemeldet, damit ihre Patienten in Zukunft (noch) besser behandelt werden können. Wenn ein Patient ein individuelles EPD angelegt hat, kann die Ärztin die Gesundheits- und Behandlungsdaten des Patienten einspeisen. Patient U. ist nun aber unschlüssig, ob er ein EPD eröffnen will und sorgt sich, dass seine Daten missbraucht werden könnten. Mit Verweis auf die Grundorientierung der Gerechtigkeit kann die Ärztin Patient U. wie folgt überzeugen:

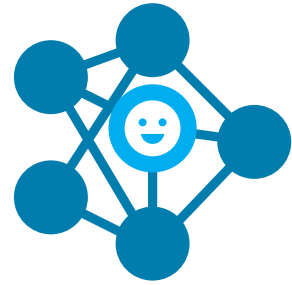
- Sie zeigt dem Patienten den individuellen Nutzen auf: Durch das EPD ist der Zugang zu den behandlungsrelevanten Dokumenten jederzeit gewährleistet und doppelte respektive unnötige Behandlungen können vermieden werden.
- Sie verdeutlicht dem Patienten, dass er selbst bestimmen kann, wer auf welche Gesundheitsdaten Zugriff hat. Nur im Notfall kann auch ohne die Einwilligung des Patienten auf das EPD zugegriffen werden. Die Daten werden dezentral nur auf Servern in der Schweiz gespeichert, wobei der Austausch der Informationen nur über eine sichere Verbindung erfolgt und dass alle Prozesse eigens zertifiziert wurden.
- Sie verweist schliesslich auch auf die Möglichkeit, dass der Patient seine Gesundheitsdaten auf einfache Weise auch der Forschung zur Verfügung stellen kann, wenn er dies wünscht.

## AUTONOMIE

Die Empfehlungen zur Grundorientierung Autonomie sollen den Nutzerinnen und Nutzern ein ausreichendes Mass informationeller Selbstbestimmung geben:

Empfehlungen

- Bieten Sie Ihren Kunden Werkzeuge, um Entscheidungen über die Zustimmung zur Datennutzung mit hoher Benutzerfreundlichkeit treffen zu können.
- Entwickeln Sie dafür intuitiv einleuchtende und gegebenenfalls durch den Nutzer wählbare Formen von informierter Zustimmung.
- Ermöglichen Sie es den Personen, ihre Zustimmung später anzupassen oder zu widerrufen.
- Erlauben Sie nach Möglichkeit die Änderung von Präferenzen über den Zweck der Datenbearbeitung.



### BEISPIEL – VERSCHIEDENE STUFEN DER ZUSTIMMUNG

Eine Web-Plattform liefert Patientengruppen Informationen über medizinische Forschung zu bestimmten Krankheiten, Möglichkeiten der sozialen Vernetzung und Self-Tracking Apps. Die Plattform sammelt, analysiert und vermarktet Benutzerdaten für kommerzielle Zwecke, aber auch für die Erstellung der personalisierten Empfehlungen, die für die Funktionalität und Attraktivität der Plattform entscheidend sind. Das Geschäftsmodell der Plattform besteht darin, sowohl anonymisierte als auch persönliche Patientendaten pharmazeutischen Unternehmen zur Verfügung zu stellen, um Patienten für die Erprobung neuer Medikamente zu rekrutieren und mit der pharmazeutischen Industrie und Universitäten zusammenzuarbeiten, um Verhaltensforschung zu betreiben (z.B. über die Einstellung der Patienten zu Therapien und deren Nebenwirkungen).

- Das Unternehmen bietet verschiedene Stufen der Einwilligung an: Nutzer können in die Verwendung ihrer Daten für Empfehlungen einwilligen, ohne dass die persönlichen Daten an die Industrie weitergegeben werden.
- Die Nutzer können ihre unterschiedlichen Einwilligungen und Präferenzen jederzeit überprüfen. Eine Zusammenfassung dazu kann leicht über das Patienten-Dashboard abgerufen werden.
- Ein Benutzer kann jederzeit entscheiden, keine Daten mehr mit der Industrie zu teilen – während bereits geteilte anonymisierte Daten nicht betroffen sind, werden persönliche Daten nicht mehr geteilt, sobald die Zustimmung elektronisch widerrufen wird.



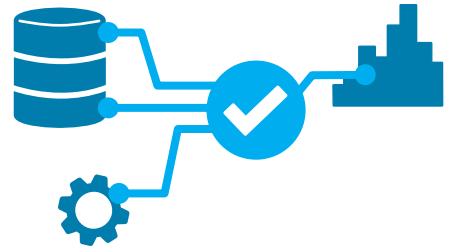
## PROZEDURALE WERTE

### KONTROLLE

Die Empfehlungen zum prozeduralen Wert Kontrolle sollen sicherstellen, dass bei der Erzeugung bzw. dem Erwerb von Daten alle relevanten Informationen bereitgestellt werden, um die Beurteilung der ethischen Nutzung der Daten zu gewährleisten:

Empfehlungen

- Dokumentieren Sie Ihre Datenquellen und jede Einschränkung bezüglich der Verwendung der Daten auf angemessene Weise.
- Wenn andere Unternehmen Ihre Daten nutzen, stellen Sie ihnen die Informationen zur Verfügung, die sie benötigen, um Datenherkunft und Nutzungsbeschränkungen zu beurteilen.



### BEISPIEL – OFFENE DATEN FÜR NEUE GESCHÄFTSMODELLE

Ein grosser Transportdienstleister beschliesst, die im eigenen Unternehmen gesammelten Daten kostenfrei und in maschinenlesbarer Form Dritten zur Sekundärnutzung zur Verfügung zu stellen (Open Data). Sie wollen damit kreative Marktteilnehmer bei der Entwicklung neuer, attraktiver Informations- und Service-Angebote unterstützen. Man erhofft sich damit die Förderung neuer Geschäftsmodelle mit innovativen Informations- oder Serviceangeboten und damit letztlich eine Steigerung der Attraktivität des Marktumfelds des Unternehmens. Das Unternehmen definiert dabei die Nutzungsbedingungen wie z.B. Pflicht zur Quellenangabe, eine Pflicht zur Aktualisierung der Rohdaten durch den Datennutzer (sofern neue Daten durch das Unternehmen zur Verfügung gestellt wurden) um zu verhindern, dass Geschäftsmodelle auf veralteten Daten beruhen und Pflichten zur Einhaltung des Datenschutzes – also insbesondere das Unterlassen jeglicher Versuche, aus den Daten deanonymisierte Informationen über die Kunden des Unternehmens zu gewinnen.

## TRANSPARENZ

Die Empfehlungen zum prozeduralen Wert Transparenz sollen sicherstellen, dass das datenschutzrechtliche Erfordernis der informierten Einwilligung auf eine für die Nutzerinnen und Nutzer möglichst einfache, klare und eingängige Weise erreicht werden kann:



### Empfehlungen

- Finden Sie einfache Wege, um zu erklären, welche Daten Sie sammeln, wie Ihre Datenschutzpolitik aussieht und wie Sie die Daten Ihrer Kunden verwenden werden (für welchen Zweck und welche Ergebnisse Sie erwarten).
- Verwenden Sie Zusammenfassungen und andere Möglichkeiten, um die zentralen Punkte Ihrer allgemeinen Geschäftsbedingungen (AGB) darzulegen.
- Arbeiten Sie mit anderen Unternehmen zusammen, um neue Standards für besser lesbare und verständliche AGBs zu schaffen.

### BEISPIEL - FOTOGRAF

Ein Fotograf bietet unter anderem Fotoshootings an. Seine allgemeinen Geschäftsbedingungen (AGB) sehen vor, dass die Fotos den Kunden digital über einen externen Online-Dienstleister zur Verfügung gestellt werden können, so dass die Kunden diese von dort herunterladen können. Der Fotograf kann die Fotos nach der Auftragsabwicklung zum Zweck der Nachbestellungen für ein Jahr weiter aufbewahren und hat zudem das Recht, die Fotos für eigene Zwecke zu verwenden, insbesondere Webseite, Plakate und ähnliches. Das Urheberrecht an den Fotos bleibt beim Fotograf. Eine Kundin wünscht nun Ausführungen über die AGBs. Der Fotograf kann diesen Wunsch wie folgt erfüllen:

- Er händigt die AGB der Kundin vor Vertragsabschluss aus. Er fasst zusammen, zu welchen Zwecken die Daten verwendet werden und verpflichtet sich, die Daten nicht zu anderen Zwecken zu verwenden.
- Er weist die Kundin darauf hin, dass er die Fotografien nach der Auftrags erledigung löschen kann, sofern sie auf eine Nachbestellung verzichten will.
- Er fragt die Kundin explizit, ob er die Fotografien für eigene Werbezwecke verwenden darf.
- Er stellt schliesslich auch sicher, dass sich der externe Online-Dienstleister an die gleichen Regeln hält, was den Umgang mit den Daten betrifft (insbesondere nicht die Daten für eigene Zwecke benutzt) und informiert die Kundin entsprechend.





## RECHENSCHAFT

Die Empfehlungen zum prozeduralen Wert Rechenschaft sollen sicherstellen, dass Sie die von Dritten gesammelten Daten auf verantwortliche Weise nutzen können:

Empfehlungen

- Wenn Ihr Produkt von Daten anderer Unternehmen abhängt, stellen Sie sicher, dass diese auf ethische Weise gesammelt wurden.
- Verzichten Sie auf die Verwendung von Daten aus nicht vertrauenswürdigen Quellen.
- Erwerben Sie nur Daten von Partnern, die bezüglich ihrer Datenerhebungspraktiken und bezüglich allen mit der Nutzung der Daten verbundenen Einschränkungen transparent sind.



### BEISPIEL - KOMMERZIELLE NUTZUNG VON SOZIALEN NETZWERKDATEN

Ein pharmazeutisches Unternehmen bezieht gegen Entgelt Daten von einer Web-Plattform, die Patientengruppen soziale Netzwerke und Möglichkeiten zur Selbstverfolgung bietet. Das Pharmaunternehmen erhält Informationen über die Praktiken der Datenerfassung der Webplattform. Werden die Daten so beschafft, dass der Zweck der Datenerhebung und die Verbindung mit der Industrie, einschliesslich der Monetarisierung der Daten, den Benutzern nicht transparent und offen kommuniziert werden? Wenn ja, wird das pharmazeutische Unternehmen keine Daten von dieser Plattform erwerben und deren Dienstleistungen nicht nutzen.

## SCHRITT 2: DATENSPEICHERUNG UND -MANAGEMENT

### ETHISCHE GRUNDORIENTIERUNGEN

#### SCHADENVERMEIDUNG

Die Empfehlungen zu Schadenvermeidung sollen sicherstellen, dass mit den gespeicherten Daten kein Schaden angerichtet werden kann.

Empfehlungen

- Implementieren Sie geeignete Massnahmen für Cybersicherheit, die kohärent auf alle Formen der Datenspeicherung (z.B. cloud-basiert oder hybrid), angewendet werden.
- Entwickeln Sie Notfallpläne, um im Fall einer Datenpanne das Ausmass des Schadens zu minimieren.
- Definieren Sie Regeln für das Löschen von Datensätzen mit dem Ziel, das Schadenrisiko eines unberechtigten Zugangs zu verringern.



#### BEISPIEL - GESUNDHEITSDATEN IN DER CLOUD

Ein Unternehmen für die Analyse von Gesundheitsdaten entwickelt massgeschneiderte Datenerfassungs- und Analysetools für grosse Krankenhäuser. Das Unternehmen stellt den Krankenhäusern einen Cloud-Service für das Hosting aller Daten zur Verfügung, so dass sich die Krankenhäuser auf die Patientenversorgung und die Behandlungsoptimierung konzentrieren können. Um den Krankenhäusern die notwendige Cybersecurity zu garantieren, implementiert das Datenanalytik-Unternehmen die höchsten Standards der Cybersecurity und verpflichtet sich den Krankenhäusern gegenüber zu ihrer Einhaltung (z.B. die Daten sind nur für ausgewählte Mitarbeiter des Software-Unternehmens mit Zwei-Faktor-Authentifizierung zugänglich und es gibt im Spital Sensibilisierungskurse für social engineering Angriffe).



## GERECHTIGKEIT

Die Empfehlungen zu Gerechtigkeit sollen dazu beitragen, dass Kundinnen und Kunden in die Lage versetzt werden, ihre Daten auch zum Wohle Dritter und des Gemeinwohls einsetzen zu können.

### Empfehlungen

- Unterstützen Sie Ihre Kunden dahingehend, dass sie das Recht auf Datenportabilität effektiv nutzen können.
- Versetzen Sie Ihre Kunden in die Lage, ihre Daten (insofern sie nicht anonymisiert wurden) auf einfache Weise zu erhalten und wiederzuverwenden, damit deren Daten in geeigneter Form auch für Anwendungen Dritter genutzt werden können.



### BEISPIEL - DATENSPENDE

Eine Firma, die personenbezogene Daten von Kunden erfasst und bearbeitet, um damit Dienstleistung zu erbringen, will es ihren Kunden leichter machen, die eigenen Daten für Zwecke zu spenden, die den Kunden sinnvoll erscheinen. Dies können beispielsweise gemeinnützige Projekt der Forschung oder auch der öffentlichen Verwaltung sein. Durch die Wiederverwendung von Daten will das Unternehmen ausserdem einen Beitrag zur digitalen Nachhaltigkeit leisten. Folgende Massnahmen könnten hier sinnvoll sein:

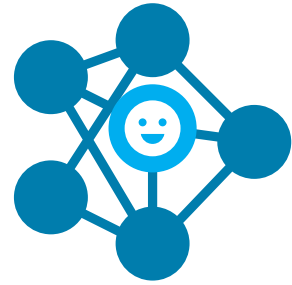
- Einrichtung eines Informationsmanagement-Systems auf der Firmen-Plattform, mit dem die Kunden der Firma ihre persönlichen Daten selbst verwalten können. Dies beinhaltet insbesondere eine benutzerfreundliche Möglichkeit, alle oder ausgewählte Daten in maschinenlesbarer Form herunterzuladen, um diese in andere Systeme zu exportieren.
- Einrichtung einer technischen Schnittstelle zu anderen Informationsmanagement-Systemen und Datentreuhand-Plattformen, so dass die Kunden der Firma ihre persönlichen Daten direkt auch anderen Dienstleistern zur Verfügung stellen können.

## AUTONOMIE

Die Empfehlungen zu Autonomie sollen sicherstellen, dass Datenspeicherung und -management keine Risiken für die Privatsphäre der Kundinnen und Kunden bilden.

### Empfehlungen

- Untersuchen Sie, ob Ihre technischen Massnahmen zum Schutz der Privatsphäre Ihrer Kundinnen und Kunden angemessen und verhältnismässig sind.
- Informieren Sie sich regelmässig über den neusten Stand entsprechender Technologien.
- Stellen Sie sicher, dass gespeicherte anonymisierte Daten nicht zur ursprünglichen Person, auf die sie sich beziehen, zurückverfolgt werden können.
- Bieten Sie Ihren Kunden unkomplizierte Möglichkeiten, die Speicherung von deren Daten einzuschränken.



### BEISPIEL – VISUALISIERUNG DIGITALER SELBSTBESTIMMUNG

Eine Firma hat die digitale Selbstbestimmung ihrer Benutzer ins Zentrum ihrer Geschäftstätigkeit gestellt und stellt ein persönliches Informationsmanagement-System und einen Marktplatz für persönliche Daten bereit. Die Plattform ermöglicht es ihren Benutzern, ihr Datenprofil oder ausgewählte Ausschnitte davon einzelnen Firmen, Verwaltungen oder auch Non-Profit-Organisationen zu einem bestimmten Preis oder auch unentgeltlich zur Verfügung stellen; die Firma partizipiert am bezahlten Preis. Die Benutzer haben jederzeit den Überblick, welche Daten zu welchem Preis an welche Datenbezüger gehen und können diese Angebote auch wieder kündigen. Der Mehrwert der Plattform für die Benutzer besteht in der Stärkung ihrer digitalen Selbstbestimmung durch folgende Elemente:

- Die Daten der Benutzer, wie z.B. Einkaufsdaten, Mobilitätsdaten oder Gesundheitsdaten werden aggregiert und visualisiert. So können die Benutzer beispielsweise überprüfen, wie gesund oder nachhaltig ihre Einkäufe waren oder wie ihr Mobilitätsverhalten in den letzten Monaten war.
- Die Daten werden verschlüsselt gespeichert und sind für die Firma selbst nicht einsehbar.
- Wenn ein Kunde Teile seines Datensatzes an Dritte weitergeben will, wird automatisch geprüft, wie hoch das Risiko ist, ob der so definierte Datensatz auf die ursprüngliche Person zurückverfolgt werden kann. Bei hohem Risiko wird der Kunde vor einer Weitergabe gewarnt, und muss ihr explizit zustimmen.



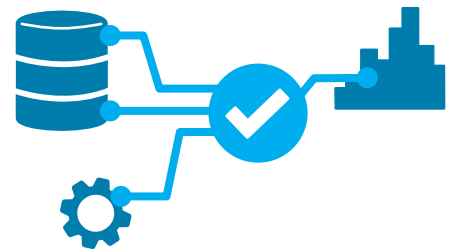
## PROZEDURALE WERTE

### KONTROLLE

Die Empfehlungen zu Kontrolle sollen sicherstellen, dass beim Management von Daten (also z.B. data cleaning, Löschen von Daten, etc.) keine für die Beurteilung der ethischen Nutzung der Daten relevanten Informationen verlorengehen.

Empfehlungen

- Dokumentieren Sie jede Änderung und Anpassung der Datenverwaltung und geben Sie die entsprechende Information an alle Nutzer der Daten (unternehmensintern und -extern) weiter; insbesondere, wenn Sie die Datenverwaltung als Dienstleistung an externe Anbieter auslagern.
- Implementieren sie technische Voraussetzungen zur Sicherstellung der Rückverfolgbarkeit (Traceability) von Änderungen an den gespeicherten Daten.
- Überlegen Sie sich im Voraus, was mit den Daten inaktiver Konten bzw. ehemaliger Kunden zu tun ist.



### BEISPIEL - NACHVOLLZIEHBARKEIT DER DATENNUTZUNG

Eine Firma verwendet soziodemographische Daten ihrer Kunden sowie Daten aus der Kundeninteraktion (Bestellverhalten, Korrespondenz, etc.), um ihre Kunden mittels datenanalytischer Modelle in verschiedene Kundengruppen einzuteilen, und um das zukünftige Verhalten von Kunden zu prognostizieren. Diese Modelle werden im Customer Relationship Management verwendet. Ausserdem werden die Resultate der modellbasierten Prognosen zum Teil mit anderen Unternehmen aus der gleichen Unternehmensgruppe geteilt. Diese Daten ändern sich ständig. Um zu jeder Zeit nachvollziehen zu können, welche Daten für die weitere Verarbeitung (Datenanalysen, Prognosemodelle) verwendet wurden, ist jeder Datensatz mit Metadaten angereichert, die beschreiben, woher die Daten kommen, zu welchem Zeitpunkt die Daten an die Analysten oder Modellbauer weitergeliefert wurden, welche Nutzungseinschränkungen mit dem Datensatz verknüpft sind, etc. Damit kann jeder einzelnen Datensatz rückverfolgt werden, der für Analysen oder Modelle verwendet wurde, und es ist gleichzeitig sichergestellt, dass die Analysten und Modellbauer die Daten in adäquater Art und Weise verwenden.

## TRANSPARENZ

Die Empfehlungen zu Transparenz sollen sicherstellen, dass Kundinnen und Kunden über Datenspeicherung und -management ausreichend informiert werden, um Vertrauen zu schaffen.

### Empfehlungen

- Verdeutlichen Sie Ihren Kundinnen und Kunden, wie ihre Daten geschützt werden und wie Zugriffe geloggt werden.
- Wenn es notwendig ist, Daten für eine sehr lange Zeit zu speichern (z.B. aus rechtlichen Gründen), erklären Sie dies Ihren Kunden zum Zeitpunkt der Datenerfassung.



### BEISPIEL - INFORMATION ÜBER VORRATSDATENSPEICHERUNG

Das Bundesgesetz betreffend die Überwachung des Post- und Fernmeldeverkehrs verpflichtet die Telekomanbieter in der Schweiz, die Metadaten der Telekommunikation ihrer Kundschaft zu speichern und sechs Monate lang aufzubewahren. Damit soll die Aufklärung von Strafdaten und die Suche von vermissten Personen erleichtert werden. Jeder Schweizer hat zudem das Recht, Einsicht in seine auf Vorrat gespeicherten Daten zu erhalten. Ein Telecom-Provider macht auf diesen Sachverhalt angemessen und detailliert aufmerksam, indem er aufzeigt, welche Daten dies sind:

- Wer wann auf welcher Website war bzw. welche Apps genutzt hat
- Mit wem wann und von wo aus kommuniziert wurde
- Wer sich wann und für welche Dauer ins Internet eingeloggt hat
- Wer wann wem ein E-Mail oder eine andere Textnachricht geschickt hat
- Wo sich der Handynutzer gerade befindet bzw. den Standortverlauf der letzten sechs Monate



## RECHENSCHAFT

Die Empfehlungen zu Rechenschaft sollen sicherstellen, dass eine ausreichende Klärung der Rollen aller am Datenmanagement beteiligten Personen stattgefunden hat.

Empfehlungen

- Definieren Sie auf der Ebene der Business-Organisation, wer an der Bewertung, Berichterstattung und Umsetzung datenethischer Standards beim Datenmanagement beteiligt ist.
- Definieren Sie klare Regeln für den Zugang zu gespeicherten Daten (wer, wann, unter welchen Bedingungen, Nachverfolgung des Zugangs, ...).



### BEISPIEL - DIE ROLLE DES CHIEF DATA OFFICER

In vielen Organisationen hat sich die Rolle des Chief Data Officers (CDO) etabliert. Der CDO beaufsichtigt die Nutzung von Daten um sicherzustellen, dass eine Organisation den grösstmöglichen Nutzen aus ihren Daten ziehen kann, dabei aber eigene Standards einhält. Dazu gehören typischerweise das gesamte Datenmanagement und die Sicherstellung der Datenqualität. Beides sind Elemente, die in der vorliegenden Empfehlung angesprochen werden. Obwohl das Amt des Chief Data Officers üblicherweise nicht aus ethischen Erwägungen heraus eingeführt wird, sondern aus Business-Überlegungen, gehört es gleichwohl zu seinen Aufgaben, datenethische Standards einzuführen und zu überwachen, wenn solche vorgesehen werden.

## ETHISCHE GRUNDORIENTIERUNGEN

### SCHADENVERMEIDUNG

Die Empfehlungen zu Schadenvermeidung sollen sicherstellen, dass im Prozess der Datenanalyse und Modellgenerierung Risikopotenziale der Modelle früh genug erkannt werden.

#### Empfehlungen

- Beurteilen Sie die vorhersehbare missbräuchliche Nutzung Ihrer Modelle.
- Vermeiden Sie im Zug der Modellentwicklung insbesondere jegliche Massnahmen, welche zur De-Anonymisierung der verwendeten Daten führen kann.
- Integrieren Sie in Ihre Modelle Kontrollmechanismen zur Überwachung und Begrenzung des potenziellen Schadens, der bei deren praktischen Verwendung entstehen könnte; dies betrifft insbesondere die Verletzung der Privatsphäre oder fehlerhafte Entscheidungen aufgrund Ungenauigkeiten im Modell.
- Erwägen Sie gegebenenfalls die Auferlegung von Grenzen für den Vertrieb des Modells.



### BEISPIEL - KI-SICHERHEITSBEWERTUNG

Einer der Hauptakteure auf dem Gebiet der KI (Deep Mind) hat einen theoretischen Ansatz für die Bewertung der Sicherheit seiner datenbasierten Produkte entwickelt. Der Ansatz besteht aus drei Elementen: (1) Spezifikation, die wir im Folgenden als Beispiel für Schritt 3 Kontrolle betrachten; (2) Robustheit: Identifizierung von Situationen, die nicht im Datensatz repräsentiert sind und zu Fehlverhalten der KI führen können (z.B. Haushaltsreinigungsroboter, der in einem Haus ohne Haustiere getestet wurde, kann in einem haustierfreundlichen Umfeld eingesetzt werden), sowie Test der Möglichkeit, kontradiktorische Eingaben zu machen; (3) Sicherung, die aus zwei Aspekten besteht: Inspektionssysteme (zusammenfassende Statistiken, automatisierte Inspektionen und Werkzeuge, die dazu dienen, KI-Entscheidungen nach dem Gebrauch erklärbar oder interpretierbar zu machen) und Durchsetzungsmechanismus (z.B. Unterbrechbarkeit, ein zuverlässiger Aus-Schalter).

Quelle: <https://deepmind.com/blog/article/specifying-ai-safety-problems>





## GERECHTIGKEIT

Die Empfehlungen zu Gerechtigkeit adressieren insbesondere das Problem der indirekten Diskriminierung aufgrund datenbasierter Modelle und Vorhersagen. Indirekte Diskriminierung bedeutet, dass eine Entscheidungsregel zwar für alle in gleicher Weise gilt, sich aber auf bestimmte Personengruppen (z.B. Frauen vs. Männer) anders auswirkt als auf andere und dies nicht gerechtfertigt ist. Dies ist insbesondere relevant für Modelle und Prognosen, aufgrund deren automatisch Entscheidungen getroffen werden, oder welche die menschliche Entscheidungsfindung unterstützen (z.B. Empfehlungen).



Empfehlungen

- Prüfen Sie, ob ein für maschinelles Lernen verwendeter Datensatz vielfältig genug ist, um Verzerrungen (Bias) aufgrund schlechter Trainingsdaten zu vermeiden.
- Erklären Sie, welche Art von Unterscheidung beabsichtigt und in Ihr Modell eingebaut ist, welche unerwünschte indirekte Diskriminierung gegebenenfalls aus Ihrem Modell beseitigt wurde, und welche Form indirekter Diskriminierung noch vorhanden ist.
- Prüfen Sie das Vorliegen von indirekter Diskriminierung mit Testdaten.
- Entwickeln Sie unternehmensintern Fachwissen über den Umgang mit indirekter Diskriminierung, die aus maschinellem Lernen resultieren kann; dies soll auch technische (Statistik, Informatik) Methoden zur Beseitigung solcher Ungerechtigkeiten beinhalten.
- Führen Sie gegebenenfalls relevante Technologien zur Vermeidung indirekter Diskriminierung ein.

### BEISPIEL - VERZERTES DIAGNOSTISCHES WERKZEUG

Ein interdisziplinäres Team von Medizinern erstellt ein Vorhersagemodell für das Ansprechen von Patienten auf eine Behandlung auf der Grundlage von Daten aus der Genomik, Epigenomik, Proteomik und anderen Bereichen. Das Modell liefert auf der Grundlage dieser Daten eine Vorhersage für eine bestimmte Behandlung, ob sie bei diesem bestimmten Patienten erfolgreich sein wird oder nicht. Die Behandlung kann das Leben des Patienten retten und hat nur geringe Nebenwirkungen. Aufgrund von Ressourcenbeschränkungen wird jedoch entschieden, die Patienten zu priorisieren, die am ehesten von der Behandlung profitieren. Es stellt sich heraus, dass 90% der Männer europäischer Abstammung, aber nur 10% der Männer afrikanischer Abstammung voraussichtlich von der Behandlung profitieren werden, so dass es ein Missverhältnis zwischen den Personen gibt, denen die Behandlung empfohlen wird. Dies wirft folgende Frage auf: Haben Schwarze eine begründete Beschwerde gegen diese Ungleichheit? Zunächst wird argumentiert, dass die ungleichen Chancen, von der Behandlung zu profitieren, diese Ungleichheit rechtfertigen. Ein Datenwissenschaftler weist jedoch darauf hin, dass es möglich ist, die Sensitivität der Vorhersage für Schwarze zu erhöhen (d.h. ein grösserer Anteil schwarzer Patienten, die geheilt werden können, wird identifiziert), mit dem Nebeneffekt, dass dies ihre Präzision verringert (d.h. einem grösseren Anteil schwarzer Patienten, die nicht geheilt werden können, wird fälschlicherweise vorhergesagt, dass sie geheilt werden können). Um diesen Kompromiss zu bewerten, wird ein Stakeholder-Panel unter Beteiligung von Patienten- und Antidiskriminierungs-Organisationen organisiert. Das Gremium kommt zu dem Schluss, dass der angepasste Algorithmus im Interesse der schwarzen Patienten ist, für die (im Gegensatz zu weissen Patienten) keine andere wirksame Behandlung zur Verfügung steht.

## AUTONOMIE

Die Empfehlungen zu Autonomie sollen dazu beitragen, dass auch Betroffene ausserhalb des Unternehmens in geeigneter Weise Einschätzungen über ethische Aspekte der Modellbildung geben können, falls diese Modelle einen Einfluss auf die Autonomie dieser Akteure haben.

### Empfehlungen

- Identifizieren Sie relevante ethische Fragen, die mit der Verwendung Ihrer datenbasierten Modelle oder Ihres aus den Daten gewonnenen Wissens verbunden sind, zusammen mit den betroffenen Anspruchsgruppen.
- Ermöglichen Sie gegebenenfalls Betroffenen oder anderen Anspruchsgruppen, Kommentare zur Modellentwicklung zu liefern; insbesondere bezüglich der in die Modellierung eingebetteten Zielvorstellungen (z.B. bezüglich des Klassifikationsziels oder der Verlustfunktion).



### BEISPIEL – WORLD CAFÉ MIT STAKEHOLDERN

Ein Schweizer Telekommunikationsunternehmen pflegt den Dialog mit ihren wichtigsten Anspruchsgruppen jeweils thematisch auf Basis seiner Nachhaltigkeitsstrategie. Unter anderem haben 25 Vertretende externer Anspruchsgruppen im Rahmen eines World Cafés mit dem Unternehmen über wesentliche Nachhaltigkeitsthemen diskutiert. Hierbei wurden drei wesentliche Themen identifiziert:

- Aufklärung über Risiken des Mobilfunkstandards G5.
- Schulung von Lehrpersonen und KMU zu Themen der Digitalisierung durch Schaffung von Digitalisierungs-Labs für KMU und Schulen.
- Schaffung vermehrter Anreize für Lieferanten zur CO<sub>2</sub>-Reduktion.

Die Teilnehmenden des World Cafés haben die Möglichkeit sehr geschätzt, sich direkt mit dem Unternehmen auszutauschen. Aufgrund der positiven Rückmeldungen beabsichtigt das Unternehmen, in Ergänzung der bewährten Stakeholderumfrage weitere Diskussionsanlässe mit Einbezug ihrer Anspruchsgruppen durchzuführen. Diese Art von Stakeholder-Dialog kann auch verwendet werden, um gezielt über datenbasierte Modelle zu diskutieren, z.B. über die Gefahr, dass diese Modelle Diskriminierung erzeugen.



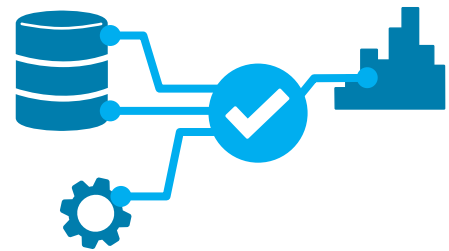
## PROZEDURALE WERTE

### KONTROLLE

Die Empfehlungen zu Kontrolle sollen sicherstellen, dass bei der Analyse von Daten ethische Aspekte systematisch dokumentiert werden.

#### Empfehlungen

- Begründen Sie die Angemessenheit der verwendeten Algorithmen im Hinblick auf den gegebenen Datensatz und die spezifische Analyse-Aufgabe; die Begründung sollte eine Dokumentation umfassen, die für die Bewertung der Risiken und der getroffenen Massnahmen zur Risikominderung im Lichte der oben genannten Grundsätze der Schadenvermeidung, Gerechtigkeit und Autonomie relevant ist.
- Stellen Sie sicher, dass jedes generierte Wissen und erzeugte Modell bis zu den ursprünglichen Daten und dem für seine Erstellung verwendeten Algorithmus zurückverfolgt werden kann.
- Wenn Ihre Analyse auf Daten und/oder Modellen anderer aufbaut, stellen Sie sicher, dass dieser Input in einer Weise erstellt wurden, die Ihren eigenen ethischen Standards entspricht.



### BEISPIEL - KONTROLLE VON KI-MODELLEN

Einer der Hauptakteure auf dem Gebiet der KI (Deep Mind) hat einen theoretischen Rahmen entwickelt, um die Bewertung der Sicherheit seiner Produkte zu leiten. Der Rahmen hat drei Elemente, wobei das erste Element, die Spezifikation, den Wert der Kontrolle implementiert. Der Konstrukteur unterscheidet die «ideale Spezifikation» (die Wünsche des Konstrukteurs), die «Entwurfsspezifikation» (z.B. die in einem reinforced-learning-Algorithmus verwendete Belohnungsfunktion) und die «aufgedeckte Spezifikation» (das tatsächliche Verhalten des Systems). Das erste Element der Kontrolle besteht darin, die «ideale Spezifikation» (welchen Zweck Sie erreichen wollen) und die «Entwurfsspezifikation» (wie diese in Code übersetzt wird) klar zu dokumentieren. Sie verwendet eine datengesteuerte Methode (z.B. inverses reinforced learning und simulierte Szenarien), um zu ermitteln, wie die aufgedeckte Spezifikation von der idealen Spezifikation abweicht, d.h. um zu beurteilen, wann die KI von den Wünschen des Erstellers abweicht.

Quelle: <https://deepmind.com/blog/article/specifying-ai-safety-problems>

Ein ähnlicher Ansatz wird in ABOUT ML definiert, der von Partnership for AI entwickelten Methode, die Transparenz durch Dokumentation beinhaltet.

Siehe: <https://www.partnershiponai.org/wp-content/uploads/2019/07/ABOUT-ML-v0-Draft-Chapter-2.pdf>

## TRANSPARENZ

Die Empfehlungen zu Transparenz sollen sicherstellen, dass sowohl die Nutzerinnen und Nutzer der Analysen und Modelle als auch allfällige Fachleute aus dem Bereich Auditing diese hinsichtlich ihrer ethisch relevanten Eigenschaften einschätzen können. Dies betrifft insbesondere die Grenzen der Modelle, wobei die entsprechende Information auf die jeweiligen Nutzer der Modelle angepasst sein sollte.



### Empfehlungen

- Erklären Sie die Logik Ihrer Modelle und die in die Modellbildung eingeflossenen Entscheidungen auf eine Weise, die die Anwender des Modells verstehen können.
- Machen Sie den Nutzern Ihrer Modelle deren Grenzen deutlich, beispielsweise Unschärfe bei Klassifikationen.
- Legen Sie den Nutzern mögliche Folgen solcher Ungenauigkeiten der Modelle dar wie beispielsweise indirekte Diskriminierung aufgrund von modellbasierten Entscheidungen.
- Dokumentieren Sie die Massnahmen, die ergriffen wurden, um Ihre ethischen Standards in Bezug auf die datenbasierte Modellierung zu erfüllen; z.B. Ihre Aktivitäten zur Beseitigung oder Minderung indirekter Diskriminierung in Ihrem Modell, falls vorhanden.
- Informieren Sie Nutzer auf angemessene Weise über diese Massnahmen; gegebenenfalls auch die Öffentlichkeit, falls die Modelle in Entscheidungskontexten verwendet werden, welche die Menschenrechte betreffen können; informieren Sie dabei zielgruppengerecht.

### BEISPIEL - AN DIE ZIELGRUPPE ANGEPASSTE INFORMATIONEN ZU EINEM KI-MODELL

Angenommen, ein Datenwissenschaftler trainiert ein maschinelles Lernmodell mit Informationen über Kreditantragsteller, um vorherzusagen, ob sie über einen Zeitraum von zwei Jahren pünktliche Zahlungen leisten werden. Die Vorhersage des Modells wird genutzt um zu entscheiden, ob ein Antragsteller für einen Kredit in Frage kommt. Drei verschiedene Zielgruppen für eine Erklärung dieses Modells können unterschieden werden:

- 1) Datenwissenschaftler, die das Verhalten des Modells als Ganzes und nicht nur in spezifischen Fällen verstehen möchten
- 2) Kreditsachbearbeiter, welche die Entscheidung verstehen müssen, warum der Antrag eines Antragstellers im Vergleich zu anderen Anträgen angenommen oder abgelehnt wird
- 3) Kunden, die verstehen möchten, warum er oder sie keinen Anspruch auf einen Kredit hat und was er oder sie tun könnte, um dieses Urteil rückgängig zu machen.

Für jeden Anspruchsberechtigten wird eine andere Art der Erklärung mit der in der AI Explainability 360 Suite verfügbaren Software ausgewählt:

- 1) eine Software, die eine Annäherung an das Produkt des Algorithmus in Form von direkt interpretierbaren, regelbasierten Modellen erzeugt, die ein globales Verständnis seines Verhaltens ermöglichen;
- 2) eine Software, die prototypische Benutzerprofile erfolgreicher Antragsteller erzeugt;
- 3) eine Software, die Beispiele berechnet, die den Kunden sagen, welche individuelle Änderung in seinem oder ihrem Profil die Entscheidung des KI-Modells verändert hätte, wenn alles andere gleichgeblieben wäre.

Quelle: Tutorials-Seite des AI Explainability 360 Toolkit, abrufbar unter <https://aix360.mybluemix.net>



## RECHENSCHAFT

Die Empfehlungen zu Rechenschaft sollen sicherstellen, dass im komplexen Prozess der Datenanalyse und Modellgenerierung die Verantwortlichkeiten für die Einhaltung ethischer Empfehlungen gewahrt werden.

Empfehlungen

- Weisen Sie den in der Modellentwicklung involvierten Personen Verantwortlichkeiten zu, um sicherzustellen, dass ethische Standards bei der Datenanalyse und Wissensgenerierung eingehalten werden.
- Stellen Sie auf der organisatorischen Ebene sicher, dass ethische Standards in der Datenanalyse und Wissensgenerierung eingehalten werden bzw. entsprechende ethische Probleme erkannt werden können.



### BEISPIEL – DATA GOVERNANCE BOARD

Ein Medienunternehmen, das vermehrt neue Geschäftsmodelle mit Daten entwickeln will, hat ein Data Governance Board eingerichtet. Dieses trifft sich mindestens vierteljährlich (wenn nötig auch ad hoc) und involviert Vertreter aller Business-Bereiche, insbesondere auch die in der Produktentwicklung aktiven Teams sowie die Rechtsabteilung. In diesem Board werden an einem Nachmittag jeweils die ethischen Fragen neuer Geschäftsideen offen und unter Anleitung einer Person diskutiert, welche sich in einer Weiterbildung verstärkt mit ethischen Themen auseinandergesetzt hat. Die Ergebnisse des Treffens werden protokolliert. Die gemeinsame «Lernerfahrung» sensibilisieren für ethische Fragen im Berufsalltag. In kritischen Fällen gelangen diese Fragen auch an die Geschäftsleitung und werden dort diskutiert. Die Aufzeichnungen der Gruppengespräche helfen dabei, das Problem möglichst präzise darzustellen, was die Geschäftsleitung für eine effiziente Behandlung des Themas schätzt.

## SCHRITT 4: PRODUKTE UND DIENSTLEISTUNGEN

### ETHISCHE GRUNDORIENTIERUNGEN

#### SCHADENVERMEIDUNG

Die Empfehlungen zu Schadenvermeidung sollen sicherstellen, dass das Missbrauchspotenzial Ihrer datenbasierten Produkte und Dienstleistungen erkannt wird.

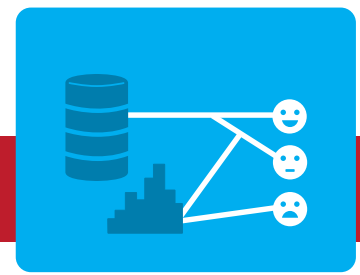
##### Empfehlungen

- Vertreiben Sie keine datenbasierten Produkte und Dienstleistungen, bevor Sie nicht untersucht haben, inwiefern unkundige Nutzer damit Schäden verursachen oder fachkundige Benutzer damit Missbrauch betreiben können.
- Prüfen Sie, inwiefern Sie erheben können, ob die von Ihnen entwickelten datenbasierten Produkte oder Dienstleistungen tatsächlich von Dritten missbraucht werden (z.B. wenn eine Software für andere Ziele als die, für die sie entworfen und getestet wurde, verwendet wird).
- Prüfen Sie, ob die Nutzung Ihres Produkts zur De-Anonymisierung von anonym vorliegenden Daten führen könnte oder zur Erhebung sensibler Daten aus ursprünglich unproblematischen persönlichen Daten verwendet werden könnte.



#### BEISPIEL – MISSBRÄUCLICHE BOT-INTERAKTION

Ein Unternehmen entwickelt unter Nutzung von künstlicher Intelligenz einen Auskunft-Bot, der analog zum früheren telefonischen Auskunft-Dienst allgemeine Fragen beantworten soll. Der Software-Bot lernt aufgrund der Interaktion mit den Nutzerinnen und Nutzer, wie zufrieden diese mit den Antworten sind; trainiert wurde er mit umfangreichen, öffentlich zugänglichen Informationen wie z.B. von Wikipedia. Auf Fragen zu kontroversen Themen gibt der Bot Antworten, die jeweils alle Aspekte der Kontroverse beinhalten, um Neutralität zu wahren. Bald wurde klar, dass der Software-Bot zunehmend auch bei faktisch gesicherten Informationen «neutrale» Antworten gibt: beispielsweise die Möglichkeit offenlässt, dass die Erde eine Scheibe sein könnte. Grund dafür war, dass Anhänger von Verschwörungstheorien in grosser Zahl Unzufriedenheit mit den Antworten des Bots ausdrückten und damit das Lernen des Software-Bots beeinflussten. Die Entwickler hatten es verpasst, systematische missbräuchliche Interaktionen mit dem Bot zu prüfen. Immerhin wurde festgestellt, dass der Missbrauch stattgefunden hat und es konnten Gegenmassnahmen eingeleitet werden.



## GERECHTIGKEIT

Die Empfehlungen zu Gerechtigkeit sollen dazu beitragen, dass die Nutzung datenbasierter Produkte und Dienstleistungen keine unerwünschten sozialen Auswirkungen wie indirekte Diskriminierung, Stigmatisierung oder Ausschluss von Personengruppen haben. Die ethische Pflicht, soziale Auswirkungen zu bewerten und anzugehen, ist proportional zu den finanziellen und technologischen Möglichkeiten eines Unternehmens sowie zum Umfang der Auswirkungen datenbasierter Produkte und Dienstleistungen auf die Gesellschaft.



### Empfehlungen

- Vermeiden Sie den Einsatz von datenbasierten Produkten und Dienstleistungen, deren indirekt diskriminierende Auswirkungen nicht untersucht und quantifiziert worden sind, wenn die Gefahr einer signifikanten indirekten Diskriminierung besteht.
- Vermeiden Sie datenbasierte Produkte und Dienstleistungen, die unbeabsichtigt ganze Personengruppen stigmatisieren oder ihren Ruf bedrohen können.
- Beurteilen Sie, ob Ihre datenbasierten Produkte und Dienstleistungen in der Anwendung vulnerable Gruppen ausschliessen, und vermeiden Sie einen solchen Ausschluss, wenn möglich.
- Überwachen Sie automatisierte Entscheidungssysteme oder Empfehlungssysteme auf unbeabsichtigte Diskriminierung, selbst wenn das Datenprodukt auf indirekte Diskriminierung getestet wurde, weil die Testdaten andere statistische Eigenschaften als die realen Daten haben können und deshalb neue Formen von Diskriminierung entstehen könnten.

### BEISPIEL - WERTKONFLIKTE IN COVID-APPS

Viele europäische Apps zur digitalen Kontaktverfolgung, wie z.B. Swiss Covid in der Schweiz, Immuni in Italien und die Corona Warn App in Deutschland, haben sich für eine dezentrale Architektur entschieden, die durch eine Google/Apple-API implementiert wird. Dies schränkt die Daten, die für weitere Analysen gesammelt werden können, erheblich ein. Dies und die dezentralisierte Architektur macht es sehr schwierig einzuschätzen, ob bestimmte soziale Gruppen ungleich von Meldungen und Falschmeldungen betroffen sind. Wenn die App zur Pflicht gemacht würde oder mit erheblichen sozialen Vorteilen verbunden wäre (z.B. die Nutzung öffentlicher Verkehrsmittel), würde die Unkenntnis darüber, welche Gruppen benachrichtigt werden, und die Ungleichheit der Falsch-Positiv-Rate bedeuten, dass wichtige Elemente der Gerechtigkeit beim Einsatz ignoriert werden. So schützt die Anwendung in der konzipierten Form zwar die Privatsphäre, macht es aber schwieriger, Gerechtigkeit zu erreichen. Wenn das Wissen über Ungerechtigkeit nicht durch eine obligatorische dezentralisierte App zur Kontaktverfolgung erzeugt werden kann, besteht ein Spannungsverhältnis zwischen Gerechtigkeit und Privatsphäre. Sollte die App also obligatorisch werden, sollten Google/Apple das Protokoll derart modifizieren, dass es die systematische Sammlung anonymisierter Daten ermöglicht, um Korrelationen zwischen der Benachrichtigungsrate und falsch-positiven Benachrichtigungen mit geschlechtsspezifischen, ethnischen oder Einkommensmerkmalen zu ermitteln und so unverhältnismässig gefährdete Bevölkerungsgruppen zu identifizieren.

## AUTONOMIE

Die Empfehlungen zu Autonomie sollen sicherstellen, dass Rückmeldungen über allfällige negative Auswirkungen Ihrer datenbasierten Produkte und Dienstleistungen ernst genommen werden.

### Empfehlungen

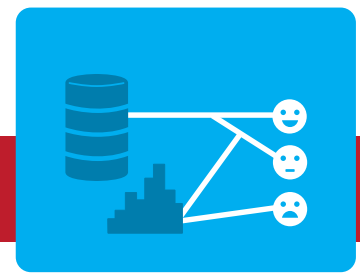
- Implementieren Sie Prozesse, die es Betroffenen ermöglichen, auf negative Ergebnisse Ihrer Datenprodukte hinzuweisen, und gehen Sie diesen Hinweisen nach.
- Verfolgen sie Experten-Debatten in den jeweiligen Anwendungskontexten ihrer Datenprodukte, die Indizien auf systematische Abweichungen in Entscheidungsprozessen basierend auf dem Wissen der jeweiligen Fachleute geben.



### BEISPIEL - BÜRGERFORUM

Deep Mind hat sich mit der (britischen) Royal Society for the encouragement of Arts, Manufactures and Commerce (RSA) zu einem Projekt zusammengetan, um ein sinnvolles öffentliches Engagement für die Auswirkungen der KI in der realen Welt zu fördern. Die RSA führte eine Bürgerjury durch, um den Einsatz von KI für die Entscheidungsfindung zu untersuchen. Dabei wurden allgemeine Fragen zur KI diskutiert, z.B. welche Art von Engagement die Öffentlichkeit in Sachen KI braucht oder was die roten Linien der Bürger sind, die vor dem Einsatz und in der Entwurfsphase relevant sind (siehe Schritt 3 – Autonomie). Um das hier beschriebene Prinzip der Autonomie zu implementieren, könnte dieselbe Methodik (Bürgerforum) als Prozess zur Erörterung negativer Aspekte bei der Einführung einer Technologie nach deren Einsatz verwendet werden. Quelle: <https://www.thersa.org/action-and-research/rsa-projects/economy-enterprise-manufacturing-folder/tech-and-society/forum-for-ethical-ai>





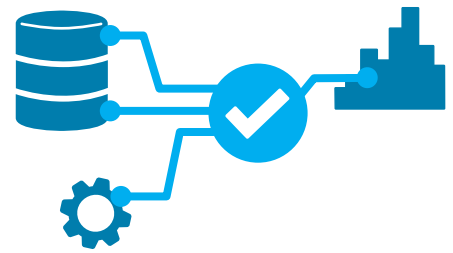
## PROZEDURALE WERTE

### KONTROLLE

Die Empfehlungen zu Kontrolle sollen auf einer allgemeinen Ebene sicherstellen, dass die unternehmenseigenen Werte sich in geeigneter Weise in den datenbasierten Produkten und Dienstleistungen niederschlagen, welche das Unternehmen auf dem Markt anbieten will.

Empfehlungen

- Bestimmen Sie die datenethischen Prioritäten für Ihr Unternehmen, die sich aus Ihren Wertverpflichtungen und Ihrer Mission als Unternehmen ergeben und leiten Sie daraus entsprechende Anforderungen an ihre datenbasierten Produkte und Dienstleistungen ab.
- Beurteilen Sie, ob ethische Standards in allen Schritten der Datenverarbeitung, die zu Ihrem Produkt oder Ihrer Dienstleistung geführt haben, auch eingehalten worden sind (unabhängig davon, ob sie innerhalb Ihres Unternehmens oder von anderen durchgeführt wurden), und ob sie mit dem Auftrag, den Werten und dem öffentlichen Image des Unternehmens in Einklang stehen.



### BEISPIEL - DURCHSETZUNGSPROBLEME VON KI-RICHTLINIEN

Die NGO AlgorithmWatch hat begonnen, Richtlinien für «ethische KI» in einem Verzeichnis zusammenzustellen. Von den mehr als 160 Richtlinien, die inzwischen in diesem AI Ethics Guidelines Global Inventory enthalten sind (2020), verfügt allerdings nur ein Bruchteil (10) über angemessene Aufsichts- und Durchsetzungsmechanismen. Durch eine regelmässige Überprüfung solcher Richtlinien wie im Fall von AlgorithmWatch und eine daran anschliessende kritische Diskussion könnte ein öffentlicher Druck entstehen, Anforderungen, die sich aus datenethischen Standards ergeben, auch kontrollierbar zu machen.

## TRANSPARENZ

Die Empfehlungen zu Transparenz sollen sicherstellen, dass Sie in angemessener Weise über die Befolgung der in diesem Kodex angegebenen Empfehlungen informieren und zur Aufrechterhaltung einer ethischen Datenpraxis beitragen.

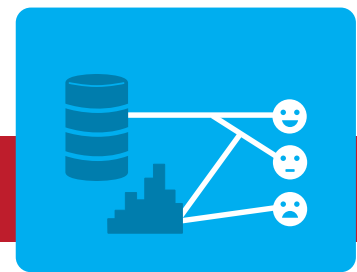
### Empfehlungen

- Kommunizieren Sie Ihre ethischen Standards und Praktiken an alle, die von den Ergebnissen Ihres Datenprodukts betroffen sein könnten.
- Wenn Ihr Unternehmen Daten, Modelle oder Datenprodukte für Dritte produziert, stellen Sie alle Informationen zur Verfügung, die zum ethischen Einsatz dieser Daten/Modelle/Produkte erforderlich sind; z.B. wenn Dritte ein Modell auf der Grundlage der von Ihnen bereitgestellten Daten trainieren, stellen Sie so viel Informationen über den Datensatz zur Verfügung, dass unbeabsichtigte Modellverzerrungen vermieden werden können.



### BEISPIEL – WEITERVERKAUF VON MODELLGENERIERTEM WISSEN

Eine Firma verwendet Daten, um über Datenanalyse neue Informationen zu erzeugen, die wiederum anderen zur Weiterverarbeitung zur Verfügung gestellt werden. Bei den Daten, welche die Firma nutzt, hat diese festgehalten, zu welchen Zwecken sie diese verwenden darf und sie hat dies in Metadaten abgelegt. In eigenen Datenethik-Richtlinien werden grundlegende Werte wie Nicht-Diskriminierung, Achtung der Persönlichkeit, Selbstbestimmung, Transparenz, Rechenschaft usw. definieren. Dabei hat sich die Firma insbesondere zur Rückverfolgbarkeit der Daten verpflichtet, d.h. dass die Herkunft und der Verwendungszweck der Daten auch innerhalb der Wertschöpfungskette bekannt bleiben. Diese Prinzipien werden in einem einfachen «Beiblatt» den jeweils verkauften Informationen beigelegt. Durch sporadische Rückfragen stellt das Unternehmen sicher, dass Drittfirmen, die ihre Produkte auf Ergebnissen und Modellen des Unternehmens aufbauen, informiert sind über die Zweckbindung der ursprünglich vom Unternehmen verwendeten Daten.



## RECHENSCHAFT

Die Empfehlungen zu Rechenschaft sollen sicherstellen, dass das Top-Management des Unternehmens in geeigneter Weise die Verantwortung für das Endprodukt übernehmen kann.

### Empfehlungen

- Zeigen Sie Führungsverantwortung, indem Sie im komplexen Prozess der Entwicklung datenbasierter Produkte und Dienstleistungen keine Verantwortungsdiffusion zulassen.
- Implementieren Sie Prozesse, um sicherzustellen, dass ethische Standards in Bezug auf die Bereitstellung des Datenprodukts eingehalten werden.
- Benennen Sie keine Sündenböcke, um die Verantwortung des Unternehmens im Fall eines ethischen Problems mit Ihren Datenprodukten abzuwälzen.
- Seien Sie sich bewusst, dass der ethische Standard Ihres Datenprodukts stark von den ethischen Standards anderer Unternehmen, mit denen Sie interagieren, beeinflusst werden könnte (z.B., weil diese die von Ihnen verwendeten Daten gesammelt haben).
- Bewerten Sie den ethischen Standard des Datenprodukts unter Berücksichtigung all dieser Einflüsse und übernehmen Sie die Verantwortung für das Endprodukt.



### BEISPIEL – CHIEF INFORMATION SECURITY OFFICER

Ein Teil der Verantwortung der Führung besteht darin, die Datensicherheit zu gewährleisten (dies ist ein Element von Schritt 2). Als Beispiel dafür kann die inzwischen bei vielen Unternehmen eingeführte Rolle des Chief Information Security Officers (CISO) dienen. Bei der Einsetzung einer solchen Funktion sollte folgendes vermieden werden:

- Mangelnde Befugnis des CISOs, so dass dieser seine Aufgabe nur formal, nicht aber effektiv, umsetzen kann.
- Installation des CISO als eines Sündenbocks, der geopfert wird, wenn ein Fehler passiert ist.

Umgekehrt können folgende Ausprägungen eines CISO zum Erfolg beitragen:

- Fokus auf Change Management und Transformation, wenn sich ein Unternehmen substantiell weiterentwickeln will
- Fokus auf operative und technische Fragen, um für den laufenden Betrieb funktionierende Lösungen zu entwickeln
- Fokus auf regulatorische Fragen und Compliance, wenn diese für das Unternehmen von grosser Wichtigkeit sind

Diese Differenzierungen sind nicht nur für den Teilbereich Information Security, sondern auch für die anderen Bereiche, die unter Schritt 1, Schritt 3 und Schritt 4 genannt werden, hilfreich.

