

## 1.1.4 Ethische Werte für die intelligente Software (inkl. KI)



■ **Stichwörter:** Ethik, Technikbewertung, Softwaregestaltung, Akzeptanz

### > Warum ist das Thema wichtig?

Die intelligente Software<sup>1</sup> mit ihren Modellen der künstlichen Intelligenz (KI) sollte ethischen Werten entsprechen, damit die Akzeptanz der Beschäftigten für die cyber-physischen Systeme (CPS)<sup>2</sup> und die 4.0-Prozesse<sup>3</sup> im Betrieb gegeben ist und die Unternehmen ihrer ethischen Verant-

wortung auch in diesen Prozessen nachkommen. Eine Berücksichtigung ethischer Werte für die Softwaregestaltung findet in einem permanenten Prozess der kritischen Überlegung und Gestaltung von legitimem, menschlich orientiertem und ökonomischem Wirtschaften gleicherma-

ßen statt.<sup>4</sup> In diesen Umsetzungshilfen werden die ethischen Werte für die Gestaltung und Programmierung der intelligenten Software (inkl. KI) formuliert, die in allen Umsetzungshilfen mit Maßnahmen konkretisiert werden.

### > Worum geht es bei dem Thema?

#### **Begriff: Ethik**

Ethik ist die Überlegung, die das menschliche Handeln aufgrund der moralischen Orientierung und den sich daraus ergebenden Ansprüchen überprüft. Die Aufgabe der Ethik besteht darin, Orientierung für die Einschätzung „Rich-

tig“ und „Falsch“ zu geben sowie für die Normen und Werte, nach denen in einer Gesellschaft oder in einem Betrieb üblicherweise darüber entschieden wird.<sup>5</sup> Die Grundannahme der Unternehmensethik ist, dass sich ethische und wirtschaftliche Zusammenhänge und Entscheidungen im

betrieblichen Alltag nicht trennen lassen; jede wirtschaftliche Entscheidung ist fast immer auch eine ethische Entscheidung.<sup>6</sup> Die grund- und menschenrechtlichen Werte bieten eine wichtige Orientierung in den 4.0-Prozessen.<sup>7</sup>

Für die Gestaltung der „traditionellen“ Technik (Arbeitsmittel, Anlagen, Handlungstechniken) gibt es Bewertungskriterien, die beispielsweise in der VDI-Richtlinie 3780 „Technikbewertung – Begriffe und Grundlagen“ beschrieben werden. Für intelligente Software mit ihren Modellen der KI gibt es kaum Vergleichbares, wenn ethische Kriterien in unterschiedlichen Zusammenhängen intensiv diskutiert werden<sup>8</sup> und auch große Konzerne die Frage der Ethik und der KI thematisieren.<sup>9</sup> Die in der VDI-Richtlinie beschriebenen Kriterien können als Orientierung

auch für die intelligente Software (inkl. KI) dienen. Auch die Hinweise der Ethikkommission zum autonomen Fahren<sup>10</sup> und des Deutschen Ethikrates zu Big Data und Gesundheit<sup>11</sup> können bei der Gestaltung der intelligenten Software mit ihrer KI herangezogen werden.

#### **Welche Werte sollten bei der Gestaltung und beim Einsatz der intelligenten Software (inkl. KI) berücksichtigt werden?**

Auf der Grundlage der beschriebenen Vorgaben sollten als ethische Kriterien die unten stehenden Wertebereiche bei

der Entwicklung der intelligenten Software (inkl. KI) in allen Anwendungsbereichen<sup>12</sup> berücksichtigt werden. Zwischen all diesen Wertebereichen bestehen mittelbare oder unmittelbare Beziehungen. Sie beeinflussen sich gegenseitig und können teilweise in der konkreten Umsetzung in Konkurrenz zueinander stehen (zum Beispiel Funktionalität oder Persönlichkeitsentfaltung und Wirtschaftlichkeit). Die einzelnen Bereiche sollten in einem möglichst ausgeglichenen Verhältnis stehen und alle berücksichtigt werden.

Diese Umsetzungshilfe gibt Experten und Interessierten Anregungen, wie Arbeit 4.0 zu gestalten ist. Die Empfehlungen sollten an die jeweilige konkrete betriebliche Situation angepasst werden.

<sup>1</sup> Intelligente Software steuert cyber-physische Systeme (CPS) und andere autonome technische Systeme (wie Messenger-Programme). Intelligente Software nutzt Modelle künstlicher Intelligenz zusammen mit anderen Basistechnologien wie zum Beispiel Algorithmen, semantischen Technologien, Data-Mining. Intelligente Software ist autonom und selbstlernend.

<sup>2</sup> Cyber-physische Systeme (CPS) verbinden und steuern als autonome technische Systeme Arbeitsmittel, Produkte, Räume, Prozesse und Menschen beinahe in Echtzeit. Die komplette oder teilweise Steuerung übernimmt intelligente Software auf Grundlage von Modellen der künstlichen Intelligenz. Genutzt werden dazu unter anderem auch Sensoren/Aktoren, Verwaltungsschalen, Plattformen/Clouds.

<sup>3</sup> Unter 4.0-Prozessen werden hier alle Arbeitsprozesse verstanden, in denen cyber-physische Systeme (CPS) oder andere autonome technische Systeme (wie Plattformen, Messenger-Programme) beteiligt sind. 4.0-Prozesse sind in den Arbeitsprozessen bisher selten vollständig, aber in Ansätzen in allen Betrieben umgesetzt.

<sup>4</sup> vgl. Ulrich 2001, S. 428

<sup>5</sup> vgl. Körtner 1999, S. 33; Rendtorff 1999, S. 154f.; Ulrich 2001, S. 43

<sup>6</sup> Küpper & Picot 1999, S. 133

<sup>7</sup> Di Fabio 2016, S. 42ff.

<sup>8</sup> vgl. u. a. Asilomar Konferenz 2017; Europäische Gruppe für Ethik der Naturwissenschaften und der neuen Technologien 2018; High-Level Expert Group on Artificial Intelligence 2018; Krüger und Lischka 2018; Rohde 2018; UNI Global Union 2017

<sup>9</sup> vgl. u. a. Deutsche Telekom 2018; Microsoft o. J.; SAP 2018

<sup>10</sup> Ethik-Kommission 2017

<sup>11</sup> Deutscher Ethikrat 2017

<sup>12</sup> Anwendungsbereiche von CPS können sein: **Insellösungen**, Teilkomponenten und Teilprozesse (zum Beispiel einzelne Arbeitsplätze, Arbeitsmittel, Teile von Anlagen, Räume, Produkte, Assistenzsysteme) und **verkettete Prozesse** und Gesamtsystemlösungen (zum Beispiel verkettete Arbeitsmittel, Wertschöpfungskette). Außerdem **geschlossene Betriebsanwendungen** (autark – zum Beispiel Edge Computing, betriebliche Cloud), **offene Anwendungen** (zum Beispiel Public Clouds, Hersteller-Plattformen).

- **Menschliche Lebensmöglichkeiten weiterentwickeln:** Entwicklung und sinnvolle Anwendung der intelligenten Software (inkl. KI) sollte die menschlichen Lebensmöglichkeiten sichern und möglichst auch verbessern.<sup>13</sup> „Der Schutz von Menschen hat Vorrang vor allen anderen Nützlichkeitsabwägungen.“<sup>14</sup>
- **Funktionsfähigkeit:** Die intelligente Software (inkl. KI) soll erstrebte Wirkungen herbeiführen, die dem „menschlichen Nutzungsbedürfnissen entsprechen“<sup>15</sup> (Brauchbarkeit). Sie soll möglichst einfach, robust, genau, eindeutig, zuverlässig und von hoher Lebensdauer sein.
- **Wirtschaftlichkeit:** Die intelligente Software (inkl. KI) soll einen bestimmten Nutzen mit möglichst geringem Aufwand herbeiführen oder mit einem bestimmten Aufwand einen möglichst hohen Nutzen erzielen. Die Wirtschaftlichkeit umfasst auch die Verbesserung der Wertschöpfung, die Verringerung der Kosten oder Steigerung des Arbeitsergebnisses, etwa über Qualitätsverbesserungen oder höhere Mengen.
- **Sicherheit:** Von der intelligenten Software dürfen keine Gefahren für Leib und Leben ausgehen und sie darf keine Personen- und Sachschäden verursachen. Dies ist vor dem Einsatz der Software sicherzustellen (wie Betriebsversagen, Missbrauchsrisiko). Bei der Risikobewertung haben sicherheitstechnische Erfordernisse den Vorrang vor wirtschaftlichen Überlegungen<sup>16</sup> (nach DIN 3100). „Die Vermeidung von Unfällen ist Leitbild.“<sup>17</sup> Muss das autonome technische System in Ausnahme- und Notsituationen entscheiden, dürfen als Entscheidungskriterien persönliche Merkmale keine Rolle spielen (Alter, Geschlecht, körperliche oder geistige Konstitution). „Eine allgemeine Programmierung auf eine Minderung der Zahl von Personenschäden kann vertretbar sein.“<sup>18</sup> Allerdings ist die intelligente Software

(inkl. KI) so zu gestalten, dass kritische Situationen gar nicht erst entstehen können. Der Arbeitsschutz muss eingehalten sein.

- **Gesundheit:** Die intelligente Software (inkl. KI) darf das „psychische und körperliche Wohlbefinden des Menschen“<sup>19</sup> nicht gefährden. Die intelligente Software (inkl. KI) ist gesundheitsgerecht gestaltet. Im Idealfall fördert sie die Gesundheit der Führungskräfte und Beschäftigten. Dies gilt sowohl für ihre eigentliche Funktion als auch für die Menschen, die sie einsetzen und verwenden. Dazu gehört dann, zum Beispiel die Kriterien der Gebrauchstauglichkeit und der Nutzungsqualität bei der Gestaltung zu berücksichtigen.<sup>20</sup> Der Gesundheitsschutz muss eingehalten sein.
- **Fairness und keine Diskriminierung:** Die intelligente Software (inkl. KI) darf keine Menschen und Menschengruppen aufgrund sichtbarer und unsichtbarer Merkmale und Dimensionen diskriminieren. Diese Merkmale und Dimensionen sind zum Beispiel Lebensalter, Geschlecht, sexuelle Orientierung/Identität, Behinderung (physische und/oder psychische Fähigkeiten), ethnische Herkunft und Nationalität, Religion und Weltanschauung. Dabei darf es weder zu intendierter Selektion noch zu nicht intendierter Diskriminierung kommen – zum Beispiel: Data-Analysen und Algorithmen können zu Ergebnissen führen, die zu einer ungewollten Diskriminierung von Personen oder Gruppen führen.<sup>21</sup>
- **Art des menschlichen Kommunikationsverhaltens als Grundlage:** Um eine effiziente, zuverlässige und sichere Kommunikation zwischen Mensch und Maschine zu ermöglichen und Überforderung zu vermeiden, sollte sich die intelligente Software (inkl. KI) stärker dem Kommunikationsverhalten des Menschen anpassen.
- **Individuelle Ressourcen unterstützen:** Die intelligente Software (inkl. KI) soll

individuelle Ressourcen und die Bewältigungskompetenz stärken. Im Idealfall fördert die intelligente Software (inkl. KI) die Persönlichkeit, das Selbstbewusstsein.

- **Transparenz und Erklärbarkeit:** Beim Einsatz von intelligenter Software (inkl. KI) sollte es möglich sein zu erkennen, wie und warum das technische System eine bestimmte Entscheidung getroffen hat und warum es in einer bestimmten Weise gehandelt hat. Erklärbarkeit und Transparenz darf nicht durch Komplexität verschleiert werden. Für die Nutzer ist Transparenz und Erklärbarkeit wichtig, da sie nur so Vertrauen in intelligente Software (inkl. KI) und Verständnis für das autonome technische System entwickeln können. Nur so können sie verstehen, was intelligente Software (inkl. KI) tut und warum.<sup>22</sup> ▶ *Siehe Umsetzungshilfe 1.1.5 Kriterien zur Erklärbarkeit der 4.0-Technologien.*
- **Handlungsträgerschaft sichtbar machen und dokumentieren:** Intelligente Software (inkl. KI) muss klar sichtbar machen, ob die Software entscheidet (Handlungsträgerschaft besitzt) oder ein Beschäftigter die Möglichkeit des Eingreifens (des „Overrulings“) besitzt und die Verantwortung behält. Übernimmt die intelligente Software (inkl. KI) die Handlungsträgerschaft „muss die Mensch/Maschine-Schnittstelle so ausgelegt werden, dass zu jedem Zeitpunkt klar geregelt und erkennbar ist, welche Zuständigkeiten auf welcher Seite liegen, insbesondere auf welcher Seite die Kontrolle liegt. Die Verteilung der Zuständigkeiten (und damit der Verantwortung) zum Beispiel im Hinblick auf Zeitpunkt und Zugriffsregelungen sollte dokumentiert und gespeichert werden. Das gilt vor allem für Übergabevorgänge zwischen Mensch und Technik.“<sup>23</sup> ▶ *Siehe Umsetzungshilfe 1.3.3 Handlungsträgerschaft im Verhältnis Mensch und intelligente Software (inkl. KI).*

<sup>13</sup> VDI-Richtlinie 3780, S. 12

<sup>14</sup> Ethik-Kommission 2017, S. 10; Europäische Gruppe für Ethik der Naturwissenschaften und der neuen Technologien 2018, S. 18

<sup>15</sup> Ethik-Kommission 2017, S. 13

<sup>16</sup> nach DIN 3100

<sup>17</sup> Ethik-Kommission 2017, S. 10

<sup>18</sup> Ethik-Kommission 2017, S. 11

<sup>19</sup> Ethik-Kommission 2017, S. 18

<sup>20</sup> vgl. DGUV Information 215-450, 2016

<sup>21</sup> Böhme und Luhn 2014, S. 46f.; Busch 2018, S. 20ff.; Deutscher Ethikrat 2017, S. 180; Krüger und Lischka 2018, S. 21

<sup>22</sup> vgl. u. a. Europäische Gruppe für Ethik der Naturwissenschaften und der neuen Technologien 2018, S. 18; UNI Global Union 2017, S. 3

<sup>23</sup> Ethik-Kommission 2017, S. 13

- **Datensouveränität:** Die intelligente Software (inkl. KI) sollte es ermöglichen, dass Menschen ein „hinreichendes Maß an Kontrolle, Souveränität und Macht über die eigenen Daten behalten oder sich zumindest auf die Wahrung ihrer Interessen durch Dritte verlassen“ können.<sup>24</sup> Die intelligente Software (inkl. KI) sollte es ermöglichen, dass der Mensch sich auf den vereinbarten Gebrauch der Daten durch Institutionen verlassen kann (organisationelle Vertraulichkeitsstandards).<sup>25</sup> Die intelligente Software (inkl. KI) ermöglicht es, eine gegebene Zustimmung zur Verwendung personenbezogener Daten zu widerrufen beziehungsweise die zurückgezogene Zustimmung softwaretechnisch umzusetzen.<sup>26</sup> Das System sollte in der Lage sein, diese Daten ganz oder teilweise zu löschen.<sup>27</sup> Der Datenschutz muss eingehalten sein.
- **Datensicherheit:** Die intelligente Software (inkl. KI) sollte die Sicherheit der Daten gegenüber Angriffen von Dritten und gegenüber Missbrauch sicherstellen.
- **Sicher im Notfall:** Die intelligente Software (inkl. KI) muss so ausgelegt sein, dass die Notwendigkeit einer abrupten Übergabe der Kontrolle an den Beschäftigten („Notstand“) praktisch ausgeschlossen ist. In Notsituationen muss intelligente Software (inkl. KI) autonom, das heißt ohne menschliche Unterstützung, in einen „sicheren Zustand“ gelangen (was unter einem sicheren Zustand zu verstehen ist, muss betriebsspezifisch näher definiert werden).<sup>28</sup> Es muss sichergestellt sein, dass es zu keinen Datenverlusten kommt und die Systeme bei Energieausfall die Grundfunktionen weiter erfüllen (zum Beispiel über ein Notaggregat)
- **Umweltqualität:** Die intelligente Software (inkl. KI) soll die Umweltqualität (Beschaffenheit der natürlichen Umgebung) fördern und nicht gefährden – das bedeutet, sparsam mit den na-

türlichen Ressourcen umgehen, Emissionen minimieren, Immissionen und Abfall vermeiden.

- **Gesellschaftsqualität:** Die intelligente Software (inkl. KI) sollte zwischenmenschliche Beziehungen und die auf diesem Zusammenwirken basierenden gesellschaftlichen Institutionen nicht gefährden.<sup>29</sup> „Nur in Sozialkontakten kann der Einzelne die Wertschätzung seiner persönlichen Eigenart und seiner persönlichen Leistungen erfahren. Deshalb sollte darauf geachtet werden, dass bestimmte Formen der Automatisierung und der neuen Medien Menschen am Arbeitsplatz oder im privaten Bereich nicht übermäßig isolieren.“<sup>30</sup>

Es besteht die schwierige Aufgabe, das soziale Zusammenleben in einer durch Technik und intelligente Software (inkl. KI) geprägten Gesellschaft und Arbeit so zu organisieren, dass „die Handlungsfreiheit der Einzelnen möglichst wenig eingeschränkt, andererseits aber auch ein notwendiges Maß an Ordnung, Stabilität und Regelmäßigkeit gewährleistet wird“.<sup>31</sup> Die Technik sollte beherrschbar bleiben, das heißt, „Technik darf sich nicht unkontrolliert und quasi eigengesetzlich entwickeln“.<sup>32</sup> Die Gesellschaft ist gefordert die ethischen Ansprüche, Kriterien und Rahmenbedingungen für die Gestaltung und den Einsatz der intelligenten Software (inkl. KI) vorzugeben, da die Unternehmen es schwer haben, alleine diese Fragen unternehmensethisch zu klären.

Die Berücksichtigung der Werte erscheint nicht nur aus grundsätzlich ethischen Gründen erforderlich, sondern sie wird auch darüber entscheiden, wie die intelligente Software (inkl. KI) gesellschaftlich sowie in jedem einzelnen Betrieb und Arbeitsprozess akzeptiert werden wird. In den „Leitlinien für den Big-Data-Einsatz“ des Bundesverbandes Informationswirtschaft, Telekommunikation und neue Medien (Bitkom) wird darauf hingewiesen,

dass sich nachhaltiger Erfolg mit Big-Data-Technologien nur dann einstellen wird, „wenn die Unternehmen hiermit auf Akzeptanz bei den Nutzern stoßen. Eine zentrale Voraussetzung dafür ist eine ethisch fundierte und untermauerte Herangehensweise“.<sup>33</sup> Den Führungskräften und Beschäftigten sollten die Sinnhaftigkeit und der Nutzen der 4.0-Prozesse vermittelt werden. „Die Akzeptanz von Big-Data-Technologien in der breiten Öffentlichkeit setzt insbesondere voraus, dass die Technologieanbieter eine ethisch fundierte und untermauerte Herangehensweise praktizieren. Ein verantwortungsvoller Umgang mit Big Data zeichnet sich dadurch aus, dass die Rechte und Interessen der Betroffenen bei der Verwendung von Big Data gewahrt werden und damit ein ethisches Korrektiv zu den nahezu unbegrenzten technischen Möglichkeiten besteht“.<sup>34</sup> Das wird – so ist anzunehmen – zunehmend eine Voraussetzung, um nachhaltig wirtschaftlich erfolgreich sein zu können.

#### Wo und wie werden diese ethischen Werte festgelegt, nach denen die intelligente Software (inkl. KI) funktioniert?

Die Werte der intelligenten Software (inkl. KI) entwickeln sich in der *Phase der Festlegung der Ziele und der Konzeption für die Software und der Programmierung*: Ziele, Konzepte und inhaltliche Anforderungen für intelligente Software (inkl. KI) werden primär von betrieblichen Akteuren (zum Beispiel von Unternehmen) definiert. Sie wird im Idealfall in Feedback-Schleifen (mit zum Beispiel IT, Fachverantwortlichen, Anwendern, Interessenvertretung) weiterentwickelt und in klassischen Lasten-/Pflichtenheften<sup>35</sup> beschrieben. Diese Ziele und Konzepte werden dann von Personen zum Beispiel mithilfe spezieller Programmiersprachen und Algorithmen, semantischer Technologien oder Ontologien und Modellen künstlicher Intelligenz programmiert. Die Präferenzen und strategischen Vorstellungen der Personen und Organisationen, die die Ziele und Konzeptionen festlegen, entscheiden über die „Werte“, nach

<sup>24</sup> Deutscher Ethikrat 2017, S. 115

<sup>25</sup> Deutscher Ethikrat 2017, S. 130

<sup>26</sup> Deutscher Ethikrat 2017, S. 160

<sup>27</sup> Deutscher Ethikrat 2017, S. 177

<sup>28</sup> Ethik-Kommission 2017, S. 13

<sup>29</sup> vgl. VDI-Richtlinie 3780, S. 20

<sup>30</sup> Ethik-Kommission 2017, S. 21

<sup>31</sup> Ethik-Kommission 2017, S. 21

<sup>32</sup> Ethik-Kommission 2017, S. 22

<sup>33</sup> Ethik-Kommission 2017, S. 71

<sup>34</sup> Bitkom 2015, S. 70

<sup>35</sup> Das Lastenheft wird in der Regel vom Auftraggeber verfasst, das Pflichtenheft vom Auftragnehmer. Das Pflichtenheft ist mit dem Auftraggeber abzustimmen.

denen die intelligente Software (inkl. KI) entscheidet und sich gegebenenfalls weiterentwickelt. Die Werte, auf deren Grundlage intelligente Software (inkl. KI) agiert, entscheiden sich damit wesentlich in der Phase der Konzeption (Lastenheft) und Programmierung der Software. Intelligente Software (inkl. KI) erhält somit ihre grundlegende Ausrichtung durch Personen und Organisationen.<sup>36</sup>

Werden fertige Tools verwendet, wird über die Werte, nach denen die intelligente Software (inkl. KI) sich entwickelt, vom Lizenzgeber der Software entschieden.

Ist die intelligente Software (inkl. KI) in Aktion, agiert sie autonom, lernt und

entwickelt sich eigenständig weiter. Wenn sich die Software in die „falsche Richtung“ entwickelt, muss die Software abgeschaltet oder verändert werden. Insofern stellen sich für die betroffenen Führungskräfte und Beschäftigten im Betrieb die Fragen: Sind die Kriterien für die Werte, nach denen die intelligente Software (inkl. KI) sich entwickelt, bekannt? Wo wurden diese Kriterien festgelegt?

Ein Beispiel, bei dem die ethischen Werte der intelligenten Software (inkl. KI) keine adäquate Berücksichtigung fanden, ist der Social-Bot von Microsoft, der im März 2016 rechtsradikal und sexistisch wurde.<sup>37</sup> Die Programmierer hatten ungewünschte Wer-

te nicht abgefangen und die intelligente Software (inkl. KI) entwickelte sich eigenständig in die „falsche“ Richtung.

Nur wenn ethische Werte in der Programmierung berücksichtigt werden, wird die intelligente Software (inkl. KI) diese Werte unterstützen und erfüllen können. Nur so kann verhindert werden, dass die intelligente Software (inkl. KI) ethische Werte im Betrieb, zum Beispiel Effizienz- und Wirtschaftlichkeitswerte, Sicherheit und Gesundheit, ignoriert. Nur so kann erreicht werden, dass Prävention integrierter Bestandteil der intelligenten Software (inkl. KI) wird.

## › Welche Chancen und Gefahren gibt es?

**Chancen:** Werden ethische Werte in der Programmierung der intelligenten Software (inkl. KI) berücksichtigt, wird diese Software am Arbeitsplatz und auch gesellschaftlich von den Menschen akzeptiert werden. Dann kann die Software

ihre Potenziale in den Arbeitsprozessen voll entfalten.

**Gefahren:** Werden ethische Werte in der Programmierung der intelligenten Software (inkl. KI) ignoriert, besteht das Risiko, dass die Software von den Men-

schen nicht akzeptiert wird und Gefahren für Wirtschaftlichkeit, Gesundheit und Sicherheit, Kommunikations- und Lebensmöglichkeiten, Persönlichkeitsentfaltung, Umwelt, Gesellschaft entstehen.

## › Welche Maßnahmen sind zu empfehlen?

Diese Maßnahmen können beispielsweise helfen, dass die verwendete intelligente Software (inkl. KI) auch ethischen Werten entspricht:

■ Es wird im Betrieb festgelegt, welche ethischen Werte die intelligente Software (inkl. KI) erfüllen sollte, wie zum Beispiel:

- Werte nach gesetzlichen Mindeststandards in den Bereichen Sicherheit und Gesundheit bei der Arbeit oder Umweltschutz.
- Darüber hinausgehende Werte wie Handlungsfreiheit, Selbstbestimmung, Erhalt sozialer Kontakte, gesundheitsgerechte Gestaltung, Datensouveränität.
- Werte, die die 4.0-Technologie<sup>38</sup> ermöglichen sollten, wie Transpa-

renz und Erklärbarkeit, Interventionsmöglichkeiten, Datenschutz, Datensicherheit.

› *Siehe auch Umsetzungshilfe 1.1.3 Unternehmensethik und intelligente Software (inkl. KI) und die oben beschriebenen Werte.*

- Alle betroffenen Personen im Betrieb werden über diese Kriterien informiert. Beschaffer und Führungskräfte werden verpflichtet, diese Kriterien bei der Anschaffung und beim Einsatz von intelligenter Software (inkl. KI) zu berücksichtigen. Zu empfehlen ist es, die ethischen Werte, die die intelligente Software (inkl. KI) im Betrieb erfüllen soll, mit den Führungskräften und Beschäftigten zu vereinbaren.
- Bei Anschaffung von smarten Dingen

(wie zum Beispiel Arbeitsmitteln, Anlagen, Arbeitsstoffen, Fahrzeugen, Einrichtungen) wird recherchiert, ob die intelligente Software (inkl. KI) dieser smarten Dinge die ethischen Anforderungen des Betriebes an die Software erfüllt. Informationen beim Anbieter/Hersteller einfordern. › *Siehe Umsetzungshilfe 2.1.5 Beschaffung digitaler Produkte.*

- Bei der Programmierung beziehungsweise der Anschaffung von intelligenter Software (inkl. KI) wird sichergestellt, dass sie die ethischen Werte des Betriebes erfüllt. Ist dies bei fertigen Tools nicht der Fall, sollte die Möglichkeit bestehen, betriebspezifische ethische Werte zu integrieren.

## Quellen und weitere Informationsmöglichkeiten:

Asilomar Konferenz (2017). *Die KI-Leitsätze von Asilomar*. <https://futureoflife.org/ai-principles-german/?cn-reloaded=1>. Zugegriffen: 03.11.2019.

Bender, G. (2007). Wechselwirkung zwischen Technik und institutionellen Strukturen

versus Technologieentwicklung als Institutionalisierungsprozess. In U. Dolata & R. Werle, R. (Hrsg.), *Gesellschaft und Macht der Technik* (S. 45–62). Frankfurt am Main, New York: Campus Verlag.

Bitkom Bundesverband Informationswirt-

schaft, Telekommunikation und neue Medien e. V. (2015). *Leitlinien für den Big-Data-Einsatz*. Berlin: Bitkom.

Böhme, R. & Luhn, S. (2014): Die Privatsphäre des Verbrauchers – ein Luxusgut? In C. Bala & K. Müller (Hrsg.) (2014), *Der glä-*

<sup>36</sup> vgl. Bender 2007, S. 46

<sup>37</sup> Steiner 2016

<sup>38</sup> 4.0-Technologie bezeichnet hier Hardware und technologische Produkte (wie Assistenzmittel/ Smartphones, Sensoren/Aktoren in smarten Arbeitsmitteln, Fahrzeugen, Produkten, Räumen etc., smarte Dienstleistungen, Apps), die von Software 4.0 (inkl. KI) ganz oder teilweise gesteuert werden.

- serne Verbraucher – Wird Datenschutz zum Verbraucherschutz? (S. 41–56). Düsseldorf: Verbraucherzentrale NRW.
- Busch, C. (2018): *Algorithmic Accountability*. Osnabrück: Universität Osnabrück.
- Deutsche Telekom (2018). *KI-Leitlinien*. Bonn: Deutsche Telekom AG.
- Deutscher Ethikrat (2017): *Big Data und Gesundheit – Datensouveränität als informationelle Freiheitsgestaltung*. Berlin: Deutscher Ethikrat.
- DGUV Information 215-450 (2016). *Softwareergonomie*. Berlin: Deutsche Gesetzliche Unfallversicherung e. V. (DGUV).
- Di Fabio, U. (2016). *Grundrechtsgeltung in digitalen Systemen*. München: Verlag C. H. Beck.
- DIN 31000:2011-05. VDE 1000:2011-05. *Allgemeine Leitsätze für das sicherheitsgerechte Gestalten von Produkten*.
- Ethik-Kommission (2017). „Automatisiertes und vernetztes Fahren“. Bericht Juni 2017. Berlin. [https://www.bmvi.de/SharedDocs/DE/Publikationen/DG/bericht-der-ethikkommission.pdf?\\_\\_blob=publicationFile](https://www.bmvi.de/SharedDocs/DE/Publikationen/DG/bericht-der-ethikkommission.pdf?__blob=publicationFile). Zugegriffen: 11.05.2018.
- Europäische Gruppe für Ethik der Naturwissenschaften und der neuen Technologien (2018). *Erklärung zu künstlicher Intelligenz, Robotik und „autonomen“ Systemen*. Brüssel: Europäische Kommission.
- High-Level Expert Group on Artificial Intelligence (2018). *Draft Ethics Guidelines for Trustworthy AI*. Brüssel: Europäische Kommission.
- Körtner, H.-J. (1999). *Evangelische Sozialethik*. Göttingen: Vandenhoeck & Ruprecht.
- Krüger, J. & Lischka, K. (2018). *Damit Maschinen den Menschen dienen Lösungsansätze, um algorithmische Prozesse in den Dienst der Gesellschaft zu stellen – Arbeitspapier*. Gütersloh: Bertelsmann Stiftung.
- Küpper, H.-U., & Picot, A. (1999). Gegenstand der Unternehmensethik. In W. Korff et al. (Hrsg.). *Handbuch der Wirtschaftsethik, Band 3: Ethik wirtschaftlichen Handelns* (S. 132–148). Gütersloh: Gütersloher Verlagshaus.
- Microsoft (o.J.). Microsoft AI-Prinzipien. <https://www.microsoft.com/de-de/ai/our-approach-to-ai>. Zugegriffen: 03.11.2018.
- Rohde, N. (2018). *Gütekriterien für algorithmische Prozesse – Eine Stärken- und Schwächenanalyse ausgewählter Forderungskataloge*. Gütersloh: Bertelsmann Stiftung.
- Rendtorff, T. (1999). Selbstverständnis und Aufgabe der Ethik. In W. Korff et al. (Hrsg.), *Handbuch der Wirtschaftsethik, Band 1: Verhältnis von Wirtschaft und Ethik* (S. 152–207). Gütersloh: Gütersloher Verlagshaus.
- SAP (2018). Die Grundsätze für Künstliche Intelligenz von SAP. <https://news.sap.com/germany/2018/09/ethische-grundsatzekuenstliche-intelligenz/>. Zugegriffen: 03.11.2018.
- Steiner, A. (2016). *Zum Nazi und Sexisten in 24 Stunden*. In Frankfurter Allgemeine Zeitung, 24.03.2016.
- Ulrich, P. (2001). *Integrative Wirtschaftsethik* (3. Aufl.). Bern, Stuttgart, Wien: Verlag Paul Haupt.
- UNI Global Union (2017). *Die 10 wichtigsten Grundsätze für ethische künstliche Intelligenz*: [http://www.thefutureworldofwork.org/media/35484/uni-global-union\\_-kuenstliche-intelligenz.pdf](http://www.thefutureworldofwork.org/media/35484/uni-global-union_-kuenstliche-intelligenz.pdf). Zugegriffen: 03.11.2019.
- VDI-Richtlinie VDI 3780 *Technikbewertung – Begriffe und Grundlagen*. September 2000.

### Zu diesem Thema könnten Sie auch folgende weitere Umsetzungshilfen interessieren:

- 1.1.1 Externe und interne Strategie in der digitalen Transformation
- 1.1.3 Unternehmensethik und intelligente Software (inkl. KI)
- 1.1.5 Kriterien zur Erklärbarkeit der 4.0-Technologien
- 1.2.1 Führung und 4.0-Prozesse
- 1.3.2 Interaktion zwischen Mensch und intelligenter Software (inkl. KI)
- 1.3.3 Handlungsträgerschaft im Verhältnis Mensch und intelligente Software (inkl. KI)
- 2.1.5 Beschaffung digitaler Produkte
- 2.3.1 Datensicherheit in 4.0-Prozessen
- 2.3.2 Datenschutz in 4.0-Prozessen



**OFFENSIVE  
MITTELSTAND**  
GUT FÜR DEUTSCHLAND

**Herausgeber:** „Offensive Mittelstand – Gut für Deutschland“ – Stiftung „Mittelstand – Gesellschaft – Verantwortung“ Kurfürsten-Anlage 62, 69115 Heidelberg, E-Mail: [info@offensive-mittelstand.de](mailto:info@offensive-mittelstand.de); Heidelberg 2019

© Stiftung „Mittelstand – Gesellschaft – Verantwortung“, 2019 Heidelberg. Gemeinsam erstellt von Verbundprojekt Prävention 4.0 durch BC GmbH Forschung, Institut für Betriebliche Gesundheitsförderung BGF GmbH, Forum Soziale Technikgestaltung, Institut für angewandte Arbeitswissenschaft e. V. – ifaa, Institut für Mittelstandsforschung Bonn – IfM Bonn, itb – Institut für Technik der Betriebsführung im Deutschen Handwerksinstitut e. V., Sozialforschungsstelle Dortmund – sfs Technische Universität Dortmund, VDSI – Verband für Sicherheit, Gesundheit und Umweltschutz bei der Arbeit e. V. – gefördert vom BMBF – Projektträger Karlsruhe