KI Prüftool BeDaX

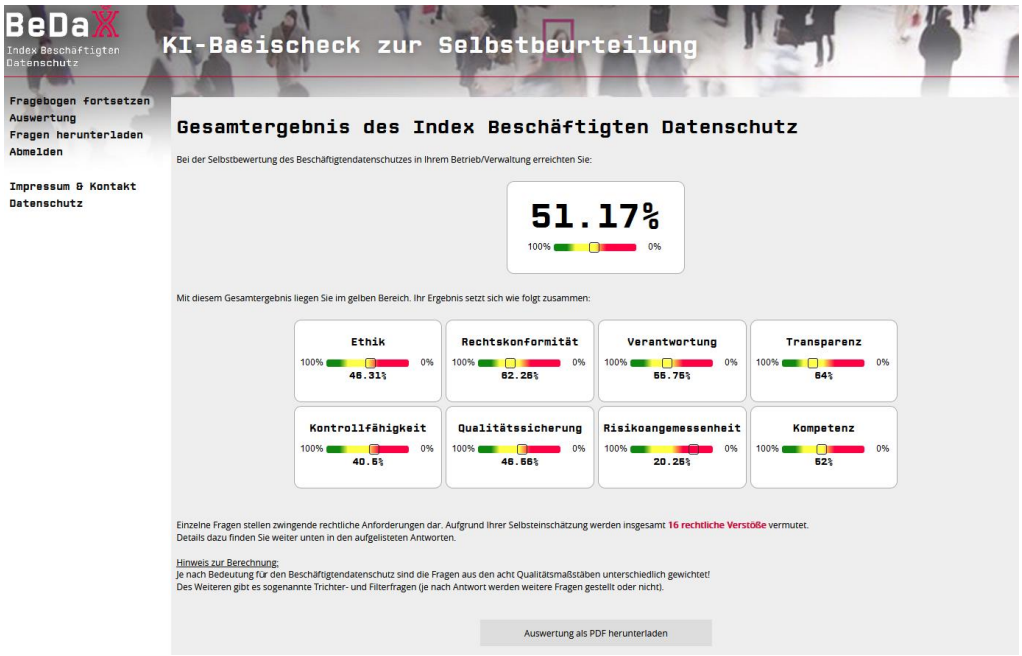
Weltweit werden immer mehr KI-Systeme in Betrieben eingesetzt, um Effizienz und Effektivität zu erhöhen, neue Geschäftsmodelle zu entwickeln, Leistungs- und Prozessqualität zu verbessern, Ressourcen einzusparen, Analytik zu verbessern und Handlungsgeschwindigkeit auszubauen. KI bietet auch in gewerblichen Einsatz immenses Potenzial, aber auch erhebliche Gefahren. Zu den Gefahren gehören die mangelnde Transparenz bei KI-gesteuerten Prozessen, Kontrollverlust, das Potenzial für Diskriminierung und das Risiko der Verletzung von Persönlichkeitsrechten. Der Schutz personenbezogener Daten ist ein Persönlichkeitsrecht das besondere Aufmerksamkeit verdient. Beim Einsatz von KI-Systemen in der Arbeitswelt sind diese so zu gestalten, dass Beschäftigte und ihre Daten geschützt, Arbeit erleichtert und die Handlungs- und Gestaltungsspieleräume der Beschäftigten erweitert werden. Interessenvertretungen (BR, PersR, MAV) spielen dabei eine zentrale Rolle.

Das Projekt BeDaX [1] (Index Beschäftigtendatenschutz) hat gemeinsam mit der Schröfers IgA GmbH und Unterstützung von ver.di ein Prüftool zur Selbstbewertung der Datenschutzaspekte bei KI-Anwendungen erarbeitet. Es ist ein Instrument zur Unterstützung der Mitbestimmungsakteure und umsichtiger betrieblicher Praktikerinnen und Praktiker. Untersucht werden die Wirkung von betrieblichen KI-Anwendungen, deren Regulierung und soziotechnischer Bedingungen entsprechend den acht BeDaX Qualitätsmaßstäben, denen 28 Faktoren zugeordnet wurden. Jeder qualitätsrelevante Faktor wird mit konkreten Fragen untersucht. Damit wird die Überprüfung des Schutzniveaus für den Beschäftigtendatenschutz bei KI-Anwendungen ermöglicht und konkrete Schutzlücken aufgezeigt.

[1] https://www.bedax.net/

Qualitätsmaßstäbe	Faktoren
Ethik	Automatisierte Entscheidungen, Mensch steuert Maschine, Profiling, Privatsphäre, unterschwellige Verhaltensbeeinflussung, personalwirksame Schlussfolgerungen, ethische Standards
Rechtskonformität	Erforderlichkeit personenbezogene Daten, Zulässigkeit der Rechtsbasis, Zweckbindung, Diskriminierungsfreiheit und Fairness, Folgenabschätzung
Verantwortung	Datenschutzsensibler Einkauf und Entwicklung, Technische Verantwortlichkeit, Fachliche Verantwortlichkeit, Verantwortung für die Betroffenenrechte, Verantwortung des/der Datenschutzbeauftragten
Transparenz	Kennzeichnung und Beschreibung, Erwartungskonformität und Nachvollziehbarkeit, Dokumentation
Kontrollfähigkeit	Aufsicht, Rückholbarkeit und Interventionsmöglichkeiten, Steuerungserfahrung und Schutz vor Überinterpretation
Qualitätssicherung	Datenqualitätsmanagement, Systemtraining, Partizipation der Beschäftigten, Evaluation
Risikoangemessenheit	Risikoeinschätzung und Klassifikation, Umgang mit untragbaren und hohen Risiken
Kompetenz	Verfügbares Fachwissen und Qualifizierungsoptionen

Die Ergebnisse der Fragen werden in dem Erhebungstool direkt nach ausfüllen der standardisierten Fragen angezeigt. Das gibt schnell wichtige Hinweise zur Ableitung von praxistauglichen, angemessenen sowie rechtskonformen Lösungen für die betrieblichen Gestaltungsarbeit.



(Beispiel der Ergebnisdarstellung KI-Basischeck zur Selbstbeurteilung, BeDaX)

Das Durcharbeiten der Fragen soll aber auch dabei helfen, mehr über die gesetzlichen Anforderungen an den Schutz der personenbezogenen Daten der Beschäftigten beim Einsatz von KI- Systemen zu lernen. Deshalb ist den Fragen jeweils ein Erklärtext vorangestellt.

(Beispiel Erklärungstext und Frage, KI-Basischeck zur Selbstbeurteilung, BeDaX)

Die spezifischen Fragen sind manchmal nicht ohne vorherige Informationsbeschaffung bei Verantwortlichen oder den Datenschutzbeauftragten beantwortbar. Das ist Absicht, es soll bei den Teilnehmerinnen und Teilnehmern an dem Selbsttest ein umfassendes Bild zum Beschäftigtendatenschutz bei KI-Anwendungen entstehen und wesentliche Elemente des Themas beleuchtet werden, die eine Schlüsselfunktion haben, die aber angesichts der Komplexität mancher Anwendungen, mangels verständlicher Erläuterungen oder der vorherrschenden Verhältnisse in manchen Betrieben einfach zu kurz kommen. Weil Rückfragen Zeit brauchen, kann der Selbsttest jederzeit unterbrochen werden und später wieder weiterbearbeitet werden.
Selbstverständlich ist jede Befragung absolut anonym und mittels einem Zugangscode geschützt.

Zu den einzelnen BeDaX Qualitätsmaßstäben wird detailliert während dem Ausfüllen des Selbsttests erläutert, was die Elemente des Qualitätsmodells, mit dem Thema KI zu tun haben:

1 Ethik

Wenn Maschinen den Menschen vermessen, bewerten und steuern, dann sind Fragen zur Ethik berührt. Beim Qualitätsmaßstab **Ethik** wird ergründet, ob bei der Anwendung von „Künstlicher Intelligenz“ Methoden des **Profiling** genutzt und **automatisierte Entscheidungen** über Personen getroffen werden. Es wird untersucht, ob und welche **personalwirksamen Schlussfolgerungen** durch Künstliche Intelligenz-Systeme, kurz KI-Systeme getroffen werden und ob diese als automatisierte Entscheidungssysteme einzuordnen sind. Es wird kritisch hinterfragt, ob der Mensch die Maschine oder die Maschine den Menschen steuert. Das hat mit Würde, aber auch mit Recht zu tun. Es werden

Indizien für Eingriffe in die **Privatsphäre** oder **unterschwellige Verhaltensbeeinflussung** der Beschäftigten ermittelt, danach gefragt, ob **KI über Karrieren entscheidet** und die Einhaltung von **ethischen Standards** thematisiert.

automatisierte Entscheidungen über Personen

Das Datenschutzrecht will die Würde und Selbstbestimmung der Beschäftigten wahren. Jeder und Jede hat deshalb nach der Europäischen Datenschutzgrundverordnung das Recht, **nicht** einer maschinellen Entscheidung unterworfen zu werden, wenn diese dem Menschen gegenüber. rechtliche Wirkung entfaltet oder in ähnlicher Weise beeinträchtigt (Art 20 DSGVO). In der Regel sind ausschließlich automatisiert getroffene Entscheidungen über Personen datenschutzrechtlich rechtswidrig. Mit der Einwilligung der Betroffenen sind solche Entscheidungsmethoden ausnahmsweise zulässig, aber selbst dann sind sie, gerade im Arbeitsverhältnis, an strenge Bedingungen gebunden. Auch eine Betriebsvereinbarung über den Einsatz derartiger Entscheidungsverfahren kann die erforderliche Einwilligung nicht ersetzen. Zu beachten ist, dass als Beschäftigte auch Bewerberinnen und Bewerber gelten. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Einsatz entsprechender Anwendungen
- Information der Beschäftigten
- verständliche Ausführungen zur Tragweite der maschinellen Entscheidungen
- Ausführungen über die Logik, mit der Entscheidungen getroffen werden
- Einwilligung, eine automatisierten Entscheidung zu akzeptieren
- Aufklärung über das Recht eine Einwilligung zu widerrufen
- Wirkung auf die Verantwortung bisheriger Entscheidungsträger

personalwirksamen Schlussfolgerungen durch KI

Die Vorgaben der Europäischen Datenschutz-Grundverordnung (DSGVO) dienen nach Art. 1 Abs. 2 und Art. 88 Abs. 1 und 2 DSGVO auch der Verwirklichung des europäischen Grundrechts auf würdige, gesunde und sichere Arbeitsbedingungen. Um diese Grundrechte im Betrieb einzulösen, gibt es in Deutschland Mitbestimmungsrechte. Sie helfen dabei, den vom Bundesverfassungsgericht formulierten Anspruch umzusetzen, wonach die Menschenwürde nicht dadurch verletzt werden darf, dass der Mensch zum Objekt herabgewürdigt wird.
Das Risiko einer Verletzung der grundrechtlich geschützten Würde besteht auch, wenn die Arbeit der Beschäftigten durch ein KI-System disponiert wird, bei dem der Mensch wie eine Maschine, wie ein Objekt behandelt wird. Werden den Beschäftigten alle Arbeitswege, Reihenfolgen der Arbeitserledigung, das Tempo und die Handgriffe oder Kommunikation durch eine KI vorgegeben, dann können Würde und Gesundheit verletzt werden. Dazu braucht das technische System nicht von Anfang an mit personenbezogenen Daten gefüttert zu werden; es genügt, wenn die Arbeit einer Person durch die KI gesteuert und kontrolliert wird. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Arbeitssteuerung durch KI
- Interventionsrechte
- Gefährdungsbeurteilungen

Profiling
Mit „Profiling“ wird die Bewertung eines Beschäftigten hinsichtlich u. a. der Arbeitsleistung, der Gesundheit, der persönlichen Vorlieben, Interessen, Zuverlässigkeit, des Verhaltens und des Aufenthaltsortes verstanden. Wird Profiling im Rahmen von automatisierten Entscheidungen betrieben, müssen Maßnahmen getroffen werden, um die Rechte, Freiheiten und Interessen der Beschäftigten zu wahren. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Einsatz entsprechender Anwendungen
- Schutzmaßnahmen
- Recht auf Einflussnahme auf und Anfechtung

Privatsphäre
Geeignete KI-Systeme können Videodaten auswerten, um Mimik und Gestik maschinell zu analysieren. Aus Audiodaten können aus der Art der Beschäftigten zu sprechen, Rückschlüsse auf deren vermeintliche Eigenarten und emotionale Verfassung gezogen werden. Es gibt zudem KI-Systeme, die aus der Analyse des Mailverkehrs Prognosen über eine Kündigungsabsicht ableiten können. Manche KI-Systeme können aus Fotos von den Beschäftigten und den von ihnen verwendeten Wörtern auf deren Einstellungen und Vorlieben schließen. Sogar Verhaltensvorhersagen können getroffen werden. Dabei sind biometrische Identifikationsverfahren, die Aussagen zum Gemütszustand oder Charakter erlauben, nach verbreiteter Rechtsauffassung unzulässig, weil sie den Menschen zum bloßen Objekt eines Verfahrens machen und ihm die Intimsphäre nehmen.
Nach europäischem Recht ist heute die Privatsphäre im Denken und Fühlen geschützt. Es ist grundsätzlich sogar untersagt, Daten z.B. zur weltanschaulichen Überzeugung und zur sexuellen Orientierung zu verarbeiten. Ein Verordnungsentwurf der EU zu Künstlicher Intelligenz vom April 2021 will die Verwendung von Emotionserkennungssystemen transparent machen und die Klassifizierung der Vertrauenswürdigkeit von Menschen durch KI-Systeme verbieten. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Auswertung von Audio und Videodaten
- Analyse elektronischer Kommunikation
- Analyse des Verhaltens von Beschäftigten in sozialen Netzwerken
- Beurteilung von Charaktereigenschaften der Beschäftigten wie Vertrauenswürdigkeit, Zuverlässigkeit, Offenheit und Loyalität durch eine KI
- innerbetrieblichen Reglementierung und Datenschutz-Folgeabschätzung

unterschwellige Verhaltensbeeinflussung

Chatbots, die Fragen von Beschäftigten beantworten, können manipulative Antworten geben; persönliche KI-Assistenten können auf Beschäftigte unterschwellig einwirken und berufliche KI-Anwendungen können Beschäftigte durch Vergleiche des persönlichen Verhaltens mit vorgegebenen Verhaltensanforderungen beeinflussen, indem sie

beispielsweise das Kommunikationsverhalten, die Produktivität oder die Pünktlichkeit vergleichend aufzeigen.
Die EU-Kommission verurteilt in ihrem Verordnungsentwurf zu „Vorschriften für Künstliche Intelligenz“ den Missbrauch von KI-Systemen für manipulative, ausbeuterische und soziale Kontrollpraktiken. Systeme, die unterschwellig beeinflussen und Personen Schaden zufügen können, sollen verboten werden. Den Beschäftigten, die mit einem KI- System kommunizieren, soll dies mitgeteilt werden. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Transparenz über den Einsatz von KI- Systemen
- Risiko der Verhaltensbeeinflussung

KI-Entscheidungen über Karrieren

Im Softwaremarkt werden KI-Systeme angeboten, um das Personalmanagement von Unternehmen zu unterstützen. Sie können Stellenanzeigen optimieren, aber auch angebotene Stellenprofile mit den Profilen von Bewerbern abgleichen. Die Systeme können Vorschläge zu personellen Entwicklungsmaßnahmen von Beschäftigten machen. Der Leistungsumfang mancher Anwendungen erlaubt es auch, Lebensläufe zu analysieren und daraus einen priorisierten Reihungsvorschlag für eine Einstellung von BewerberInnen zu erarbeiten.
Der Verordnungsentwurf der EU-Kommission zu „Vorschriften für Künstliche Intelligenz“ hält KI-Anwendungen dann für sehr risikoreich, wenn diese Entscheidungen über den Zugang zu beruflicher Bildung, Bewertung von Bewerbern, Beförderung oder Kündigung treffen. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Einsatz entsprechender Anwendungen
- Schlussfolgerungen für die Entwicklung oder Nutzung der Qualifikation

Ethik- Leitlinien

Immer mehr Unternehmen betonen ihre Orientierung auf Humanität, Offenheit, soziale Verantwortung und Diversität, indem sie sich entsprechende Leitlinien und Selbstverpflichtungen auferlegen und diese veröffentlichen. Eine Berichterstattung zu Aspekten des Umweltschutzes, der sozialen Verantwortung, den Arbeitnehmerbelangen und der verantwortungsbewussten Unternehmensführung fordern die EU, mit ihrer Richtlinie zur Corporate Social Responsibility (CSR) zumindest bei Unternehmen mit mehr als 500 Beschäftigten, der deutsche Corporate Governance Kodex und immer mehr Investoren. Manche Unternehmen bekennen sich auch zu eigenen Ethik-Leitlinien für den Einsatz von Künstlicher Intelligenz. Selbstbestimmung des Menschen, Diskriminierungsfreiheit und Gerechtigkeit: das sind einige Werte, die die Enquetekommission des deutschen Bundestages zur Orientierung für den Einsatz von KI-Systemen empfiehlt. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Existenz von Ethik- Leitlinien

- Vorgaben zu zur Einhaltung europäischer Grundrechte
- Controlling der Umsetzung der Leitlinien

2. Rechtskonformität

Rechtsfragen sind im Grunde bei den meisten Qualitätsanforderungen zum Einsatz von KI-Systemen berührt. Zum Qualitätsfaktor Rechtskonformität konzentrieren sich die Erhebungen deshalb auf die **Erforderlichkeit, personenbezogene Daten zu benutzen**, die **Zulässigkeit der Rechtsbasis** und auf die **Diskriminierungsfreiheit**. Nach der Durchführung der **Folgenabschätzung** wird gefragt.

Erfordernis, personenbezogene Daten zu benutzen

Das europäische Datenschutzrecht verlangt vom Arbeitgeber eine Risikominimierung. Auch KI-Anwendungen sollen durch datenschutzfreundliche Voreinstellungen vorrausschauend schon so geplant werden, dass die betroffenen Personen geschützt werden. Der Einsatz von personenbezogenen Daten soll beschränkt werden, beispielsweise, indem diese Daten verschlüsselt, anonymisiert oder pseudonymisiert werden. Es muss überprüft werden, ob sie für die vorgesehenen Zwecke überhaupt erforderlich sind. Die Prüfung der Erforderlichkeit kann ergeben, dass die Verarbeitung vollständig anonymer Daten zur Erreichung des legitimen Zwecks ausreicht. Überprüft werden sollte auch, ob überhaupt KI-Systeme erforderlich sind, bei denen manchmal schwer nachvollziehbar ist, wie sie über welche Kombination personenbezogener Daten zu Schlussfolgerungen kommen. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Prüfung des Bedarfs an personenbezogenen Daten
- Schlüssigkeit des Verzichts auf die Anonymisierung, Pseudonymisierung und Verschlüsselung
- Protokollierungsdaten zur Systemadministration
- Sperrung nicht freigegebener Systemteile

Zulässigkeit der Rechtsbasis

Die Verarbeitung personenbezogener Daten in KI-Systemen ohne Rechtsgrundlage ist unzulässig. Zulässig ist die Verwendung von Beschäftigtendaten durch den Arbeitgeber u.a., wenn eine Kollektivvereinbarung oder ein Gesetz dies zulässt oder die Betroffenen eingewilligt haben, weil sie dadurch einen rechtlichen oder wirtschaftlichen Vorteil haben. Dann gelten aber Prinzipien wie Fairness, Transparenz, Zweckbindung der Daten und deren Beschränkung auf das notwendige Maß. Die Daten müssen ausreichend geschützt sein und der Arbeitgeber ist in der Pflicht, sowohl die Rechtmäßigkeit als auch die Einhaltung der Prinzipien nachzuweisen. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Art der Rechtsgrundlage
- Umstände der der Einwilligung

- Aufklärung über Widerrufsrechte
- Realisierung der Informationspflichten
- Verfahren der Zweckänderung

Diskriminierungsfreiheit

Erfahrungsgemäß treffen Menschen als Entscheider mitunter auch diskriminierende Entscheidungen. Das Allgemeine Gleichstellungsgesetz (AGG) soll vor derartigen Diskriminierungen und vor sachlich nicht gerechtfertigter Ungleichbehandlung schützen. Schon das Grundgesetz enthält einen Gleichbehandlungsgrundsatz (Art. 3 GG). Auch das Datenschutzrecht will Gleichheit und Diversität am Arbeitsplatz gewährleisten. Je nachdem, wie KI entwickelt und verwendet wird, hat sie das Potenzial, Verzerrungen zu schaffen, zu verstärken, aber andererseits auch, sie zu vermeiden. Glücklicherweise lassen sich diskriminierende Algorithmen schneller umprogrammieren als etwa rassistische Personen. Es braucht dafür aber trotzdem Initiative und Kenntnis von verzerrenden Algorithmen oder unfairen Praktiken und deren Ursachen.
Bedeutsam ist es, bei den technischen Systemen die Datenauswahl und den Zuverlässigkeitsanspruch eines Systems, die Methode des maschinellen Lernens, das angestrebte Qualitäts- und Fairnessmaß sowie die Trainingsmethode des Systems, den Autonomiegrad maschineller Entscheidungen, die Anwendungsfelder, potenzielle Wirkungen und Evaluationsszenarien zu hinterfragen. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Risiko der Entscheidungssysteme für Verzerrungen und Diskriminierungen
- Verfolgter Qualitätsanspruch
- Integrität der Ursprungsdaten
- Methoden des Trainings

Folgenabschätzung

Das Datenschutzrecht verlangt vor dem Einsatz risikoreicher neuer Technologien eine Abschätzung der Folgen für die Rechte und Freiheiten natürlicher Personen. Der Bundesbeauftragte für Datenschutz und Informationsfreiheit hält eine Folgenabschätzung für den „Einsatz von Künstlicher Intelligenz zur Verarbeitung personenbezogener Daten zur Steuerung der Interaktion mit den Betroffenen oder zur Bewertung persönlicher Aspekte der betroffenen Personen“ für erforderlich.
Eine Folgenabschätzung muss u.a. eine Bewertung der Risiken z.B. für Datenschutz, die Würde der Beschäftigten und die Diskriminierungsfreiheit umfassen, die Schutzmaßnahmen beschreiben und den Nachweis der Gesetzeskonformität der Anwendung erbringen. Einsatzzwecke und Verantwortliche müssen benannt werden. Die Notwendigkeit und Verhältnismäßigkeit der Datenverarbeitung, muss bewertet werden. Es ist sinnvoll, daneben auch Folgen für die Gesundheit, die Arbeitsbedingungen und die Beschäftigtenqualifizierung zu untersuchen, zumal das Arbeitsschutzrecht danach verlangt und das Datenschutzrecht

auch den sicheren, gesunden und würdigen Arbeitsbedingungen verpflichtet ist. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Partizipation der Beschäftigten
- Transparenz der Ergebnisse
- Inhalte der Folgenabschätzung

3. Verantwortung

Wenn ihr Hund jemanden beißt, sind Sie als Hundehalter verantwortlich. Wer KI betreibt, wird sich auch nicht mit der Autonomie der Maschine herausreden können, wenn etwas passiert. Verantwortungsbewusstsein der Arbeitgeber und daraus abgeleitet ein organisiertes Verantwortungsmanagement fordert u.a. die Europäische Datenschutz-Grundverordnung. Verantwortlichkeiten müssen erkannt, angenommen und wahrgenommen werden. Es geht dabei um rechtliche, technische und administrative Verantwortung, um Verantwortung für den Schutz der Persönlichkeitsrechte sowie Verantwortung für die maschinellen Schlussfolgerungen.
Zum Qualitätsfaktor Verantwortung wird hinterfragt, inwieweit schon bei **Einkauf und Entwicklung** des KI-Systems auf angemessenen Datenschutz hingearbeitet wird. Untersucht wird, welche Aufgabe an die **technisch für das KI-System Zuständigen** übertragen wurden und wie die **Verantwortung für die maschinellen Entscheidungen** wahrgenommen wird. Die **Zuständigkeit für die Betroffenenrechte** wird thematisiert, ebenso die Zusammenarbeit mit dem / der **Datenschutzbeauftragten**

Einkauf und Entwicklung des KI-Systems
Datenschutz durch Technikgestaltung und durch datenschutzfreundliche Voreinstellungen – dazu verpflichtet die Europäische Datenschutz-Grundverordnung alle Arbeitgeber in Europa. Damit muss bei der Systemauswahl und -entwicklung begonnen werden und die Zuständigen für Einkauf und Systementwicklung sind in einer besonderen Pflicht, die besser klar beschrieben ist. Die Normen für Sicherheit einer Informationsverarbeitung verlangen darüber hinaus nach einer vorausschauenden Risikobeurteilung. Risikodimension und Risikoeintrittswahrscheinlichkeit muss auch nach dem Datenschutzrecht eingeschätzt werden.

Ohne frühzeitiges Handeln hätte jeder Arbeitgeber auch ein Problem, wenn er später seinen Informationspflichten gegenüber den Beschäftigten nachkommen will. Das Datenschutzrecht verlangt danach, dass den betroffenen Beschäftigten bei automatisieren Entscheidungssystemen Informationen über die involvierte Logik, die Tragweite und die Auswirkungen gegeben werden. Der Arbeitgeber wird verpflichtet, darauf hinzuarbeiten, dass er diese Informationen geben kann. Wird bei Systemplanung und -auswahl nicht danach gefragt, wird es schwierig, rechtmäßig zu handeln. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Leitlinien für den Einkauf oder die Entwicklung
- Information von Entscheidungsträgern und Betroffenen
- Erhebung von Referenzerfahrungen

technisch Zuständigen

Wer über die Zwecke und die Mittel der Verarbeitung personenbezogener Daten entscheidet, gilt nach dem Datenschutzrecht als Verantwortliche Stelle. In Unternehmen ist das die Geschäftsführung oder der Vorstand. Üblicherweise werden die Aufgaben für die technische Einführung und den technischen Betrieb von KI- Systemen an die Unternehmenseinheit delegiert, die ohnehin für die Informationsverarbeitungssysteme zuständig ist. Für IT-Zuständige ist es gut zu wissen, was diese Delegation für Folgen haben kann.
Das Datenschutzrecht macht detaillierte Vorgaben zur Technik von Informationsverarbeitungssystemen, also auch für KI-Systeme. Der Verstoß gegen das Recht kann sehr teuer werden. Wenn Unberechtigte Zugriff auf die Daten haben, diese manipuliert oder zweckwidrig eingesetzt werden, ist dies ein Rechtsverstoß. Rechtswidrig ist auch eine unzulängliche Folgenabschätzung oder unzureichende technisch- organisatorische Schutzmaßnahmen. Das sollten nicht nur die Geschäftsführung wissen. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Information über datenschutzrechtliche Pflichten
- Risikoverständnis
- Aufklärung über mögliche Sanktionsfolgen

fachliche Verantwortung

Fachlich bleiben Menschen verantwortlich, auch wenn lernende Maschinen Entscheidungen treffen. Fachlich Zuständige formulieren Einsatzzwecke und Ziele der KI-Systeme. Sie geben Qualitätsanforderungen und Lernziele vor. Sie bleiben auch für die Ergebnisse maschinellen Handelns verantwortlich.
Für automatisierte Entscheidungen über Personen gibt das Datenschutzrecht vor, dass die Verantwortlichen Eingriffsmöglichkeiten haben, sich für die Interessen der betroffenen Personen interessieren und sich mit einem Widerspruch eines Beschäftigten gegen den Einsatz auseinandersetzen müssen. Nicht ohne Grund warnt selbst der Deutsche Bundestag vor unklaren Verantwortlichkeiten, intransparenten Mechanismen der KI und uneindeutigen Absichten. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Abgrenzung der Zuständigkeiten
- Dokumentation der Pflichten
- Verständnis von Betroffenenrechten
- Verständnis von Sanktionsfolgen

Zuständigkeit für die Betroffenenrechte

Das Datenschutzrecht verpflichtet die Arbeitgeber bei der Verarbeitung personenbezogener Personaldaten, die Beschäftigten darüber präzise, transparent, verständlich, leicht zugänglich und in einfacher Sprache zu informieren. Die Beschäftigten haben Einwirkungs- und Anfechtungsrechte bei automatisierten Entscheidungen. Sie können eine einmal gegebene Einwilligung zur Verarbeitung ihrer Daten jederzeit widerrufen. Der Arbeitgeber soll sich für die Interessen der Beschäftigten zum Schutz ihrer Daten interessieren, und diese sollen bei der Folgenabschätzung zu Wort kommen.

Es braucht in den Betrieben zuständige Stellen für die Umsetzung dieser Rechte der Betroffenen und Pflichten der Arbeitgeber. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Transparenz der Zuständigkeitszuordnung
- Erreichbarkeit für die Betroffenen
- Bearbeitungsregeln für Interventionen

Verantwortung des/der Datenschutzbeauftragten

Betriebliche Datenschutzbeauftragte beraten die Verantwortlichen hinsichtlich deren Pflichten nach dem Datenschutzrecht, sollen die Einhaltung der Vorschriften überwachen und auf Anfrage bei der Durchführung einer Folgeabschätzung unterstützen. Dafür brauchen sie Kenntnisse über KI-Systeme und ausreichend Ressourcen, um ihre Aufgaben wahrzunehmen. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Fachkunde
- Ressourcen
- Zusammenarbeit mit Mitbestimmungsakteuren

4. Transparenz

Wenn wir nicht wissen, womit wir es zu tun haben, entstehen Ängste und wird Einflussnahme schwierig. Nicht verwunderlich, dass deshalb die meisten Expertengremien, die sich mit KI beschäftigen, Transparenz, Erklärbarkeit und Nachvollziehbarkeit von KI-Systemen zum Qualitätsmerkmal erheben. Auch das Datenschutzrecht verlangt, dass gegenüber den betroffenen Personen die Datenverarbeitung transparent und nachvollziehbar gemacht werden muss. Es legt Wert auf Verständlichkeit der Information und auch darauf, dass die Logik nach der lernende Maschinen Schlussfolgerungen ziehen, verstanden werden kann.
Zum Qualitätsfaktor Transparenz wird untersucht, inwieweit **Systeme künstlicher Intelligenz gekennzeichnet** werden, wie **nachvollziehbar** sie **beschrieben** und **dokumentiert** werden.

Hinterfragt wird auch, ob **maschinelle Entscheidungen erklärbar** sind und nach welchen Methoden die **Erwartungskonformität des Systems beurteilt** wird.

Güte der Transparenz von KI- Systemen

Um die Rechte und Pflichten nach dem Datenschutzrecht erfüllen zu können, werden Informationen darüber benötigt, ob das KI-System, oder die KI-Elemente in einem IT-System, für eine automatisierte Entscheidungsfindung herangezogen werden. Der Transparenzanspruch der Datenschutz-Grundverordnung macht es erforderlich, dass ein KI-System als solches zu kennzeichnen ist und dass präzise, verständlich, leicht zugänglich, klar und in einfacher Sprache darüber informiert wird. Bei automatisierten Entscheidungen muss über die Entscheidungslogik, die angestrebten Auswirkungen und die Tragweite der Entscheidung informiert werden. Werden die Daten nicht direkt bei den Beschäftigten erhoben, müssen diese über die Datenquellen informiert werden. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Transparenz der Wirkungen maschineller Entscheidungen
- Güte der Information der Betroffenen
- Erhebungsquelle der personenbezogenen Daten
- Einschätzung der Integrität der Daten

Nachvollzierbarkeit und Erklärbarkeit

Ein IT-System, also auch ein KI-System, gilt nach Industrienormen dann als erwartungskonform, wenn es aus dem Kontext heraus vorhersehbaren Benutzerbelangen sowie allgemein anerkannten Konventionen entspricht. Erwartungen sollten fixiert sein und transparent gemacht werden, bevor ein KI-System zum Einsatz kommt. Erst dann ist es möglich zu beurteilen, ob das System den Erwartungen entspricht, die geplanten Zwecke erreicht und alle Quelldaten erforderlich sind. Auch um zu beurteilen, ob sich eine lernende Maschine in eine unerwünschte Richtung entwickelt und Schlussfolgerungen zieht, die im Kontext nicht passen, sind Informationen über die angestrebten Zwecke erforderlich. Danach verlangt auch das Datenschutzrecht.
Um der Informationspflicht gegenüber den Beschäftigten nachzukommen, genügt es nach Auffassung der Datenschutzaufsichtsbehörden nicht, das Ergebnis automatisierter Entscheidungen zu vermitteln. Darüber hinaus muss die Nachvollziehbarkeit im Hinblick auf die Prozesse und das Zustandekommen von Entscheidungen gewährleistet sein. Das führt zu Anforderungen an die Transparenz. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Fixierung der Erwartungen an das System
- Transparenz der Vorgaben

Dokumentation
Betriebe müssen nachweisen können, dass sie den Anforderungen des Datenschutzrechts nachkommen. Sie sind gehalten, ein Verzeichnis der Verarbeitungstätigkeiten zu erstellen, und sie müssen die Personen, deren Daten sie verarbeiten, aussagekräftig über die

Datenverarbeitung informieren. Eine Dokumentation eines KI-Systems ist dafür notwendig, auch um im laufenden Betrieb noch beurteilen zu können, ob das lernende System noch „in der Spur läuft“ und nichts Falsches lernt.
Die Europäische Kommission will anspruchsvolle Anforderungen an die Dokumentation von Hochrisikosystemen stellen. Neben der Logik des KI-Systems, seiner Architektur, der Trainingsmethode und Risikoklassifikation sollen Datenquellen, Kennzeichnungsverfahren und Verfahren menschlicher Aufsicht beschrieben werden. Detaillierte Informationen werden über die Überwachung, Funktionsweise und Kontrolle, insbesondere hinsichtlich seiner Fähigkeiten, Leistungsgrenzen und Genauigkeit verlangt. Die Risiken für die Grundrechte und Grundfreiheiten sollen bezeichnet werden. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Verzeichnis der Verarbeitungstätigkeiten
- Leitlinie zur Systemdokumentation
- risikoadäquate Dokumentationstiefe
- Zugang zur Systemdokumentationen
- Güte der Systemdokumentationen
- Realisierung der Nachweispflicht

5. Kontrollfähigkeit

Menschliche Aufsicht und Kontrolle zählt zu den wesentlichsten Faktoren, die Vertrauen in KI-Anwendungen entstehen lassen. Das ist nicht verwunderlich, da KI-Systeme in ihrer Entscheidungsfindung nicht immer als nachvollziehbar, achtsam und fehlerfrei gelten können und auch nicht sicher ist, dass lernende Maschinen das Richtige dazulernen. Die EU-Kommission bringt zum Ausdruck: „Die menschliche Aufsicht dient der Verhinderung oder Minimierung der Risiken für die Gesundheit, die Sicherheit oder die Grundrechte“ beim Einsatz von KI-Systemen. Um Aufsicht und Kontrolle aber ausüben zu können, braucht es die Fähigkeiten und die Befugnisse dazu. Es braucht Handlungsmotivation, Eingriffsmöglichkeiten, alternative Arbeitskonzepte, Alarmierungssysteme und die Möglichkeit, maschinelle Entscheidungen zu revidieren.
Zum Qualitätsfaktor ‚Kontrollfähigkeit‘ wird untersucht, inwieweit zum Einsatz von Systemen Künstlicher Intelligenz **Interventionsmöglichkeiten in Form von alternativen Arbeitsprozessen** existieren und **Not-Aus-Prozeduren** eingerichtet sind. Es wird nach Maßnahmen gefragt, um **Missbrauchspotenziale und Fehlanwendungen zu erkennen und einzudämmen** und Optionen der **Rückholbarkeit (Revision)** von Systementscheidungen untersucht. Hinterfragt werden Erfahrungen mit **Überinterpretation** von Systementscheidungen und **unzureichende Steuerungsmöglichkeiten** werden thematisiert.

Interventionsmöglichkeiten, Missbrauchsschutz und alternative Arbeitsprozesse

Lernende Maschinen sollen dazulernen und sich fortentwickeln. Damit sie nicht aus der Spur laufen, will die Europäische Kommission zumindest hochriskante KI-Systeme nur zulassen, wenn diese Eingriffsmöglichkeiten in den Systembetrieb bieten und über eine Stopptaste oder ähnliche Funktionen verfügen. Fehlanwendungen sollen rechtzeitig erkannt und

behoben werden.
Nach dem Datenschutzrecht haben Personen, die von automatisierten Entscheidungen betroffen sind, sogar das Recht, dass verantwortliche Menschen in den Verarbeitungsprozess eingreifen, wenn etwas schiefläuft. Der Arbeitgeber muss Garantien für die Rechtmäßigkeit der Datenverarbeitung anbieten und hat eine Nachweispflicht für geeignete technisch-organisatorische Schutzmaßnahmen. Er würde damit gegen das Gesetz verstoßen, wenn er keine ausreichende Kontroll-, Aufsichts- und Steuerungsfunktionen für den Systembetrieb vorsieht. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Alarmierungsfunktionen
- Hinweisgeberfunktionen
- Missbrauchsschutz
- Not-Aus Mechanismen
- Alternativkonzepte
- Revisionsmöglichkeiten

Vermeidung von Überinterpretationen

KI- Systeme arbeiten nur mit den Daten, die sie bekommen. Sie vergleichen sachlich soll und ist, sie rechnen mit Wahrscheinlichkeiten, sie ziehen Schlussfolgerungen und treffen Entscheidungen zweckrational. Sie erzeugen keine Wahrheiten. Ihnen fehlt jegliches Mitgefühl, sie haben keine menschlichen Werte. Hoffnungen, Traditionen und Emotionen können sie auswerten, aber nicht empfinden. Weil uns das Wesen des Menschen wichtig ist, braucht es beim Einsatz von KI- Systemen im Betrieb Kenntnis darüber, was lernende Maschinen können und was nicht. Es braucht menschliche Aufsicht und Kontrolle, ein Verständnis von den Funktionsmechanismen der KI, es braucht Interventionsmöglichkeiten und Sicherheitsmechanismen.
Das Datenschutzrecht verlangt geeignete technisch-organisatorische Schutzmaßnahmen und Interventionsmöglichkeiten bei automatisierten Entscheidungssystemen. Die EU-Kommission will, dass sich die Verantwortlichen der menschlichen Neigung bewusst sind, übermäßiges Vertrauen in KI-Systeme zu entwickeln. Bei Daten, die von Strafverfolgungsbehörden verarbeitet werden, muss unterschieden werden, ob es sich um Tatsachen oder Einschätzungen handelt. Es ist vor diesem Hintergrund sinnvoll, auch sensible Auswertungen von KI- Systemen als maschinelle Interpretation zu kennzeichnen. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Bedeutung maschineller Schlussfolgerungen
- Widerspruchshäufigkeit
- Eingriffshäufigkeit im Testbetrieb
- Hinweise von Zuständigen
- Kennzeichnung maschineller Schlussfolgerungen

6.Qualitätssicherung

Betriebliches Qualitätsmanagement zielt u.a. darauf, gute Arbeitsergebnisse und Prozessqualität zu erhalten und auszubauen. Dafür werden Qualitätsanforderungen gesetzt und Einfluss auf Arbeitsprozesse, Organisation, Kommunikation und Qualifikation genommen. Normen, Standardisierungen und Revisionsmodelle werden genutzt, um das betriebliche Handeln auf Qualität auszurichten.
Mit den Fragen zur Qualitätssicherung bei der Einführung von KI- Systemen soll zunächst die **Eignung der Daten** überprüft werden, die das KI-System nutzt. Die Methode des **Trainings des Systems** wird untersucht. Die **Partizipation** und die **Nutzung der Expertise** von den beteiligten Akteuren im Betrieb wird hinterfragt. Die Antworten sollen Erfahrungen mit **Evaluationsmechanismen** und **Hinweisgebersystemen** im laufenden Systembetrieb zutage fördern.

Eignung der Daten

„Datenqualitätsmanagement ist im Zusammenhang mit maschinellem Lernen ein vorrangiges Thema“ – mahnt die Deutsche Normungsroadmap. Den Qualitätsanspruch an die Ursprungs- und Trainingsdaten bringt das europäische Datenschutzrecht zum Ausdruck. Die deutschen Datenschutzaufsichtsbehörden präzisieren in ihrer „Hambacher Erklärung“ die Anforderungen an die Eignung der Daten, ebenso wie die EU- Kommission in ihrem Verordnungsentwurf zu KI.
Die Daten müssen demnach mit dem ursprünglichen Einsatzzweck vereinbar, erheblich für die Nutzung durch die KI, repräsentativ, sachlich richtig, aktuell, manipulationssicher, vollständig und verfügbar sein. Die Verarbeitung personenbezogener Daten soll nicht zu Verzerrungen und Diskriminierungen führen und muss rechtmäßig und nachvollziehbar sein. Als integre Daten werden vereinfachend jene Daten bezeichnet, die den Qualitätsanforderungen entsprechen. Die betroffenen Beschäftigten müssen über die Quelle dieser Daten informiert werden. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Eignung der Eingangsdaten
- Prüfung der Integrität der Daten im Training und Test
- Offenlegung der Datenquellen

Training du Test des Systems

Damit aus „Kennen“ „Können“ wird, haben wir für das Lernen der Menschen Tests für die Feststellung des Lernerfolgs sowie Trainings zur Einübung des Erlernten organisiert. Es braucht Wiederholung und Korrektur, bis wir mit kontinuierlicher Qualität das Gelernte auch erfolgreich und kulturverträglich umsetzen können. Lernende Maschinen haben ähnliche Abhängigkeiten. Erst Tests und Training machen sie schlau. Gutes Training hilft, falsches Training schadet.
Die Enquetekommission KI des Deutschen Bundestages fordert, Trainings- und Testprozesse zu dokumentieren, und stellt fest: „Für die Güte und die Nachvollziehbarkeit algorithmischer

Entscheidungen sind die Verfügbarkeit, aber auch die Qualität, Integrität und Transparenz von Trainingsdaten und Datenmerkmalen, die der Entscheidung zugrunde liegen, von entscheidender Bedeutung.“ Die Europäische Kommission will Betriebe verpflichten, zumindest für hochriskante KI- Anwendungen Trainingsmethoden zu beschreiben und Datensätze auf Eignung und Trainingsverfahren auf Verzerrungsfreiheit zu überprüfen.
Das Datenschutzrecht unterscheidet nicht zwischen Datenverarbeitung für das Training und regulärem KI-Einsatz. Für personenbezogene Daten gelten die allgemeingültigen Grundsätze: Rechtmäßigkeit, Zweckbindung, Transparenz, Angemessenheit, Notwendigkeit, Richtigkeit, Integrität und Vertraulichkeit. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Durchführung und Legitimation des Trainings
- Trainingsziele
- Trainingsergebnisse

Partizipation

Die Bundesregierung will, dass die Erwerbstätigen bei der Entwicklung von KI-Anwendungen in den Mittelpunkt gestellt werden. Das geht nicht ohne deren Beteiligung. Das Strategiepapier der Regierung stellt fest: „Betriebliche Mitbestimmung und eine frühzeitige Einbindung der Betriebsräte stärken das Vertrauen und die Akzeptanz der Beschäftigten bei der Einführung und der Anwendung von KI. Dies ist Voraussetzung für eine positive Haltung zu KI allgemein sowie eine erfolgreiche Implementierung von KI-Anwendungen auf betrieblicher Ebene.“
Das Datenschutzrecht gibt den Beschäftigten die Möglichkeit, in der Folgenabschätzung zu Wort zu kommen. Sie müssen vom Arbeitgeber über die Verarbeitung ihrer personenbezogenen Daten informiert werden und haben bei automatisierten Entscheidungen das Recht, ihren Standpunkt darzulegen und Entscheidungen anzufechten. Beschäftigte müssen die Möglichkeit haben, auf nicht regelgerechte Anwendung von KI-Systemen aufmerksam zu machen. Sie werden am besten frühzeitig in den Prozess der Systemeinführung und Evaluation einbezogen. Dann können aus Betroffenen Beteiligte werden. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Referenzen zu Beschäftigteneinschätzungen
- Erhebung der prioritären Qualitätsindikatoren der Betroffenen
- Beteiligung im Training
- Einbeziehung der Standpunkte der Beschäftigten in die Risikoeinschätzung
- Gelegenheit um Nützlichkeit, Erwartungskonformität und Verhältnismäßigkeit des Systems zu beurteilen
- Anzahl der zurückgezogenen Einwilligungen

Evaluationsmechanismen und Hinweisgebersysteme

„Voraussetzung für einen akzeptierten KI-Einsatz in der Arbeitswelt sind partizipative, dialogische Einführungs-, Nutzungs- und Evaluationsprozesse, die bei der Festsetzung der

Ziele beginnen, eine Abschätzung der Folgen für Arbeitnehmerinnen und Arbeitnehmer anschließen und bei der Überprüfung regelgerechter Anwendung enden," schreibt der Deutsche Bundestag. Die Parlamentarier fordern die stete Evaluation, Revision und erforderlichenfalls Anpassung, Re-Design oder Beendigung des KI-Einsatzes.
Die deutsche Normungsroadmap hält „die kontinuierliche Bewertung von Leistungs- und Sicherheitsmetriken sowie die Bestimmung angemessener Reaktionen auf Zwischenfälle und die Etablierung geeigneter Gegenmaßnahmen für erforderlich". Die Überprüfungen sollen auch zutage fördern, ob menschliche Entscheidungsträger einen Kontrollverlust erleiden und ob ein blindes Vertrauen in die Schlussfolgerungen der KI entsteht. Auch die EU-Kommission plant, das Risikomanagement für riskante Anwendungen auf den gesamten Lebenszyklus des KI-Systems auszurichten.
Das Datenschutzrecht verlangt danach technisch-organisatorische Schutzmaßnamen erforderlichenfalls zu überprüfen, ebenso die Folgenabschätzung. Es besteht im Betrieb also das Erfordernis, mit einem systematischen und dokumentierten Verfahren festzustellen, ob Handlungsbedarf besteht. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Hinweise hinsichtlich einer nicht regelgerechten Verwendung eines KI-Systems
- Transparenz von Whistleblower-Verfahren
- Anzeichen für Kontrollverlust und blindem Vertrauen
- Evaluationsprinzipien

7. Risikoangemessenheit

Die Enquetekommission KI des Deutschen Bundestages, die Datenethikkommission der Bundesregierung, die Europäische Kommission und die Deutsche Normungsroadmap empfehlen einen risikobasierten Ansatz für den Umgang mit KI-Systemen. Auch Risiken für die Rechte und Freiheiten von Beschäftigten sollen untersucht werden. Differenziert nach dem Ausmaß des Risikos sollen angemessene Maßnahmen zum Abbau der Risiken ergriffen werden. Risiken sollen eingeschätzt, klassifiziert, gemindert und transparent gemacht werden. KI-Systeme haben durch ihre unterschiedlichen Funktionen, Zweck und Kontexte aber eine unterschiedliche Risikorelevanz. Es macht einen gewaltigen Unterschied, ob die KI beim Suchen oder Übersetzen hilft oder ob sie Personalentscheidungen trifft. Deshalb braucht es Sortierkriterien, um eine KI-Anwendung der richtigen Risikostufe zuzuordnen, der angemessene Schutzmaßnahmen hinterlegt sind.

Risikomanagement ist dabei nichts Neues. Schließlich fragen auch Aufsichtsräte danach und im deutschen Corporate Governance Kodex steht: „Im Lagebericht sollen die wesentlichen Merkmale des gesamten internen Kontrollsystems und des Risikomanagementsystems beschrieben werden und soll zur Angemessenheit und Wirksamkeit dieser Systeme Stellung genommen werden." Auch im Aktienrecht und im GmbH-Recht und im Genossenschaftsrecht ist vorgesehen, im Rahmen von Compliance-Maßnahmen und zur Erfüllung von Sorgfaltspflichten der Unternehmensleitung ein Risikomanagement zu implementieren (§§ 91 Abs. 2, 93 AktG; § 43 Abs. 1 GmbHG; § 34 Abs. 1 GenG).
Mit den Fragen zur Risikoangemessenheit werden die **Verfahren der Risikoeinschätzung** und **Risikoklassifikation** untersucht. Die **Realisierung eines Risikomanagementsystems** für KI-

Systeme wird zum Thema gemacht. Nach den **Regulierungsstandards für unterschiedliche Risikostufen** wird gefragt, und es wird überprüft, ob im Betrieb KI-Anwendungen eingesetzt werden, die die Europäische Kommission **verbieten** will oder als **Hochrisiko-Anwendung** betrachtet.

Verfahren der Risikoeinschätzung, Risikoklassifikation und Risikomanagement

Das Datenschutzrecht fordert, das entsprechend der Risikodimension und Eintrittswahrscheinlichkeit des Risikos einer Datenverarbeitung an der Risikobewältigung gearbeitet wird. Diese Größen muss man im Betrieb erst mal kennen. Der Verordnungsentwurf KI der Europäischen Kommission fordert mindestens für hoch riskante KI-Anwendungen, dass ein Risikomanagementsystem eingerichtet, angewandt, dokumentiert und aufrechterhalten wird. Risiken sollen beseitigt, mindestens verringert, kontrolliert und NutzerInnen zur Anwendung der Managementmaßnahmen geschult werden. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Einschätzung zu den Risikoquellen
- Art von Risiken
- Beurteilung des Risikoausmaßes
- Beurteilung der Eintrittswahrscheinlichkeit des Risikos
- Zuordnung zu Kritikalitätsstufen
- Entscheidungskriterien für die Risikoklassifikation
- Risikoadäquate Regulierung

Hochrisiko-Anwendungen und Anwendungsverbote

Die Europäische Kommission hat im April 2021 einen Verordnungsentwurf zur Regulierung von KI- Systemen der Öffentlichkeit vorgestellt. Inzwischen befindet sich der Entwurf mit Änderungsvorschlägen des Parlaments im sog. Trilog, also in der Abstimmung zwischen Kommission, Parlament und Ministerrat. Darin werden KI-Systeme benannt, die in ganz Europa verboten werden sollen, und es wird eine erste Liste von Systemen publiziert, die die EU dem Hochrisikobereich zuordnen und besonders streng regulieren will. Für die betriebliche Gestaltungsarbeit ist es sinnvoll zu überprüfen, ob derartige Systeme im Betrieb eingesetzt werden oder ihr Einsatz geplant ist. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Anwendungen, die die EU verbieten will
- Einsatzbereiche von Hochrisikoanwendungen

8. Kompetenz

Wer KI-Systemen betreibt, verantwortet, kontrolliert oder nutzt, sollte über die dafür erforderlichen Kompetenzen verfügen.
Mit den Fragen zum Qualitätsmaßstab Kompetenz‘ wird um eine Einschätzung zur Qualifikation der **Verantwortlichen** gebeten und es wird **die Sachkunde der**

Ansprechpartner der Arbeitgeber für die Interessenvertretungen betrachtet. Selbstkritisch wird die **eigene Gestaltungskompetenz** eingeschätzt; nach dem Zugang zu **Weiterbildungsangeboten** und **Sachverständigen** wird gefragt.

Fachwissen und Zugang zu Qualifikation und Sachverständigen

Der Deutsche Bundestag hat mit der Beschlussfassung des Abschlussberichtes der Enquetekommission KI die Bedeutung der Kompetenz für den Einsatz einer vertrauenswürdigen KI betont. Beschäftigten soll Nutzungskompetenz vermittelt, Arbeitnehmervertretungen soll Beurteilungs- und Gestaltungskompetenz zugänglich gemacht werden. Arbeitgeber sollen Weiterbildung zu KI anbieten. Die EU-Kommission will mindestens für Hochrisikosysteme menschliche Aufsicht vorschreiben, die in der Lage ist, Risiken für Gesundheit, Sicherheit und Grundrechte zu minimieren. Die Aufsicht soll das System vollständig verstehen und richtig interpretieren können.
Für die Aufgaben der Datenschutzbeauftragten verlangt schon das Datenschutzrecht, die Funktion auf Grundlage der beruflichen Qualifikation und insbesondere des Fachwissens zu besetzen. Folgende Aspekte werden zu diesem Qualitätsfaktor hinterfragt:

- Vorbereitung der fachlich und technisch für das KI-System Verantwortlichen
- Sachkunde der Verantwortlichen Ansprechpersonen
- eigene Gestaltungskompetenz
- Angebote für Qualifizierungsmaßnahmen
- Zugang zu KI-Sachverständigen

Gewichtung der Antworten im Selbsttest

Bei der Gewichtung der einzelnen Qualitätsmaßstäbe zu Datenschutz und künstlicher Intelligenz ist neben der **Rechtskonformität** ein **zentraler Maßstab die Ethik.** Die Kernfrage ist, ob die Maschine den Menschen oder der Mensch die Maschine steuert. Eine wichtige Grundlage zur Einflussnahme auf die ethischen und rechtlichen Aspekte stellt die **Transparenz** dar. Die Einschätzung, inwieweit das KI-System **verständlich** und die maschinellen Schlussfolgerungen **nachvollziehba**r sind, ist bedeutsam für menschliches Vertrauen. Da KI-Systeme in ihrer Entscheidungsfindung nicht immer fehlerfrei arbeiten stellt die **Kontrollfähigkeit** einen weiteren wichtigen Faktor dar. Kontrollfähigkeit korrespondiert mit der **Risikoeinschätzung**. Die Untersuchung von Aspekten der **Verantwortung,** der **Kompetenz** und der **Qualitätssicherung** ergänzen und vervollständigen den Basischeck.

Die Gewichtung der Antworten auf die Fragen zu den Qualitätsmaßstäben:

	Qualitätsmaßstäbe	%
1	Ethik	15
2	Rechtskonformität	25
3	Verantwortung	8
4	Transparenz	16
5	Kontrollfähigkeit	10
6	Qualitätssicherung	8
7	Risikoangemessenheit	10
8	Kompetenz	8
	Summe	100

Warnhinweise

Auch die Ergebnisse des KI- Selbsttests werden übersichtlich entsprechend dem o.g. Schema angezeigt und können damit dem innerbetrieblichen Dialog zugänglich gemacht werden. Darüber hinaus gibt das System aber den Personen, die sich dem Selbsttest stellen aber auch Warnhinweise, falls deren Antworten darauf hindeuten, dass Kernelemente des Datenschutzes nicht erfüllt werden und dies möglicherweise sehr kostenwirksame Sanktionsfolgen haben kann. Hier einige Beispiele zu entsprechenden Warnhinweisen:

- Bei ausschließlich automatisierten Personalentscheidungen und maschinellen Personaldispositionen haben die Beschäftigten einen Rechtsanspruch darauf, ihren Standpunkt deutlich zu machen und intervenieren zu können. Wenn eine KI die Arbeit disponiert und der betroffene Mensch nicht eingreifen kann, steuert die Maschine den Menschen und macht ihn zum Objekt maschineller Disposition. Dies ist rechtlich unzulässig. Wir empfehlen Ihnen, auf einfach zu realisierende Interventionsrechte der Beschäftigten im Betrieb zu drängen und Wert darauf zu legen, dass wesentliche personalwirksame Entscheidungen durch Menschen getroffen werden müssen
- Vorsicht: KI-Systeme, mit den Schlussfolgerungen für die Karrieren von BewerberInnen oder Beschäftigten gezogen werden, sind nach Dem Entwurf der europäischen KI- Verordnung Hochrisikoanwendungen. Wenn diese in ihrem Betrieb eingesetzt, müssen zumindest die hohen Regulierungsanforderungen erfüllt werden,

die der Verordnungsentwurf vorsieht. Wir empfehlen einen Blick in den Gesetzentwurf.

- Nach Art. 5 Abs. 1 der Datenschutzgrundverordnung gilt das Prinzip der Datenminimierung. Die Datenschutzbeauftragten des Bundes und der Länder haben in ihrer Hambacher Erklärung darauf hingewiesen, dass beim Einsatz von KI-Systemen zunächst zu prüfen ist, ob das System überhaupt mit personenbezogenen Daten arbeiten muss. Diesem Anspruch wird offensichtlich in ihrem Betrieb nicht entsprochen. Wir empfehlen spätestens beim Training des KI-Systems zu untersuchen, ob auf personenbezogene Daten verzichtet werden kann und trotzdem der Zweck des Systems erfüllt werden kann.
- Weder in herkömmliche Datenverarbeitungssysteme noch in KI-Systeme dürfen personenbezogene Daten eingebracht werden, wenn es dafür keine Rechtsbasis gibt. Eine Datenverarbeitung ohne Rechtsbasis ist rechtswidrig. Wir empfehlen beim Arbeitgeber darauf zu drängen, dass die Rechtsbasis für die Verarbeitung personenbezogener Daten durch das jeweilige KI-System präzise benannt wird, um überhaupt beurteilen zu können, ob der Einsatz des KI-Systems zulässig ist.
- Nach ihrer Einschätzung nimmt sich die in ihrem Heimatbetrieb praktizierte Folgenabschätzung nur unzureichend den für Arbeitnehmerinnen und Arbeitnehmern wesentlichen Fragen an. Eine Folgenabschätzung verlangt das Datenschutzrecht aber auch das die Arbeitsstättenverordnung und das Betriebssicherheitsgesetz. Eine Gefährdungsbeurteilung und eine vorausschauende Folgenabschätzung für die Arbeitsbedingungen der Beschäftigten und auch für deren Persönlichkeitsrechte wesentlich. Dazu zählen nach Art. 1 Abs. DSGVO die Würde der Beschäftigten, gesundheitsförderliche Arbeitsbedingungen, Privatsphäre, Datenschutz und Diskriminierungsfreiheit.
  Es kann Sanktionen für Betriebe nach sich ziehen, wenn diese Rechte nicht beachtet werden. Werden mögliche Sanktionsfolgen ignoriert, wird es den Betrieb überraschen, wie hoch Verstöße gegen die europäische Datenschutzgrundverordnung oder die europäische KI-Verordnung geahndet werden können. Mängel in der Folgenabschätzung wurden von europäischen Datenschutzbehörden schon mit Bußgeldern über mehrere Millionen € bestraft.
  Wir empfehlen eine umfassende Folgeabschätzung, die sich nicht nur auf den Schutz personenbezogener Daten ausrichtet und die Anforderungen europäischen Rechts in den Blick nimmt. Auch die Würde der Beschäftigten leitet sich aus dem Art. 1 Abs. 2 der europäischen Datenschutzgrundverordnung ab.
- Art. 13 der europäischen Datenschutzgrundverordnung verlangt danach, bei auto-automatisierten Entscheidungsystemen, die personalrelevante Entscheidungen treffen können, den betroffenen Beschäftigten die involvierte Logik zu vermitteln. Die Information soll nach Art.12 der DSGVO präzise, verständlich, leicht zugänglich, klar und in einfacher Sprache verfasst sein. Der Umstand, dass den Beschäftigten ihres Betriebes die Information darüber verweigert wird, dass diese mit einem KI System zusammenarbeiten, verstößt die guten Sitten, die europäische KI-Verordnung und bei der Verwendung personenbezogener Daten, auch gegen das europäische

Datenschutzrecht. Wir empfehlen den Einsatz von KI-Systemen, deren Zweck und Einsatzgebiete im Betrieb transparent zu machen und umso mehr Information zur Verfügung zu stellen, je risikoreicher das jeweilige KI-System klassifiziert wurde.

- In ihrem Betrieb werde nach ihrer Einschätzung allenfalls wirtschaftliche und Datenschutzrisiken von KI- Systemen untersucht. Fragen danach, inwieweit Anwendungen den Grundrechten nach der europäischen Grundrechtscharta entsprechen, werden vernachlässigt. Wir empfehlen den Arbeitgeber darauf hinzuweisen, dass er nach Art. 1 Abs. 2 und 88 DSGVO und § 26 BDSG die Verpflichtung hat, auch die Risiken zu betrachten die sich aus dem Einsatz von KI-Systemen für die Grundrechte und Grundfreiheiten natürlicher Personen ergeben (u.a. Würde, Gleichheit, Diversität, Gesundheit, Sicherheit am Arbeitsplatz und der mit der Beschäftigung zusammenhängenden Rechte und Leistungen)
- Offensichtlich werden in ihrem Betrieb KI Einwendungen eingesetzt oder sollen eingesetzt werden, die die europäische Kommission verbieten will. Wir empfehlen den Arbeitgeber darauf aufmerksam zu machen und gemeinsam mit dem Datenschutzbeauftragten darauf hinzuweisen, dass der Einsatz von rechtswidrigen Systemen erhebliche Sanktionsfolgen nach sich ziehen kann.

Zugang zum Selbsttest

Wer Interesse an der Durchführung eines KI- Basis-Selbsttests hat, kann einen Zugangscode beim Projektleiter Karl-Heinz (Charly) Brandl (brandl@input-consulting.de) erhalten. Alle Informationen dazu finden Sie auch unter www.bedax.net.

Auf dieser Homepage (www.bedax.net) finden sie auch ein umfangreiches Informationsangebot zum Beschäftigtendatenschutz, das laufend ausgebaut und erweitert wird. Es richtet sich an Mitbestimmungsakteure und vermittelt neben Grundlagen des Beschäftigtendatenschutzes, aktuelle Rechtsentwicklungen, Empfehlungen der Aufsichtsbehörden, nützliche Links und hilfreiche Praxisbeispiele.