

# „Die Sicherheit hängt davon ab, wie künstliche Intelligenz gestaltet ist“

## Key Facts

- Die DGUV Test Projektgruppe Künstliche Intelligenz (KI) veröffentlichte die DGUV Test Information 05 mit Grundsätzen zur sicherheitstechnischen Bewertung von KI
- Entwicklungsbegleitende Prüfungen helfen die Sicherheit von KI-Systemen zu verbessern
- Die sichere Gestaltung von KI-Systemen bietet Chancen für die Zukunft

## Autoren

- ➔ **Hermann Haase**
- ➔ **Dr. André Steimers**

**Künstliche Intelligenz (KI) kann die Sicherheit und Gesundheit im beruflichen Umfeld verbessern. Allerdings sind die Bewertung und Prüfung der Sicherheit von KI-Systemen eine große Herausforderung. Hermann Haase und Dr. André Steimers aus der DGUV Test Projektgruppe Künstliche Intelligenz beantworten Fragen zu Nutzen und Voraussetzungen der sicheren Gestaltung von KI.**

**Der erste Grundsatz der DGUV Test Information 05 zu künstlicher Intelligenz betont die große Bedeutung von klassischen Technologien. Das klingt widersprüchlich.**

**STEIMERS:** Am Anfang eines Projektes sollte immer die Frage stehen: Ist künstliche Intelligenz oder kurz KI überhaupt notwendig? Das mag trivial klingen. Aber KI ist „en vogue“. Forschungsvorhaben erhalten eher eine Förderung, wenn sie in Verbindung mit KI stehen. Die Finanzierung eines Forschungsprojektes oder die Faszination, die mit dem Thema verbunden ist, sollte jedoch nicht der Grund für ein KI-System sein.

Klassische Technologien sind für viele Vorhaben ausreichend oder sogar besser geeignet. In diesen Fällen sollte man auf den Einsatz von KI verzichten, denn insbesondere Verfahren des Deep Learnings, also der Einsatz tiefer neuronaler Netze, sind sehr komplex und bringen viele Schwierigkeiten mit sich. Ausschlaggebend sollte

einzig und allein der Nutzen des Systems sein. Daher der Grundsatz: künstliche Intelligenz nur, wenn es wirklich die beste Lösung ist und einen eindeutigen Vorteil bringt.

**Wie sicher ist künstliche Intelligenz denn?**

**HAASE:** Systeme auf der Basis künstlicher Intelligenz bieten zumindest nach derzeitigem Stand keine hundertprozentige Sicherheit. Dennoch können sie wertvolle Ergänzungen darstellen, die zu einer erhöhten Sicherheit führen. Nehmen wir ein Beispiel: In einem Lkw ist ein Kamera-Monitor-System eingebaut, kombiniert mit einer KI-Funktion zur Personenerkennung. Es warnt mit einem akustischen Signal, wenn sich eine Person hinter dem Lkw befindet, und erhöht somit die Wahrscheinlichkeit, dass Personen rechtzeitig erkannt werden. Aber: Wird die Fahrerin oder der Fahrer unaufmerksam, weil sie oder er sich vollständig auf das akustische Signal verlässt, kann dies zu Gefährdun-

gen führen. Denn die Personenerkennung funktioniert nicht immer.

Ein anderes aktuelles Beispiel aus der Arbeit der DGUV Test Prüf- und Zertifizierungsstelle Holz: Hier können durch künstliche Intelligenz gesteuerte Automatismen, die eine Hand erkennen und eine Säge bei Gefahr stoppen, eine deutlich erhöhte Sicherheit bieten – auch ohne hundertprozentigen Schutz wird sie die Anzahl der Unfälle deutlich verringern können.

Die Sicherheit von KI-Systemen hängt außerdem stark davon ab, wie die KI gestaltet ist: Der Einsatz mehrerer, sich gegenseitig kontrollierender Systeme bietet Verbesserungspotenzial. Ob das Umfeld eines Fahrzeuges personenfrei ist, kann von KI-gestützten und nicht KI-gestützten Systemen gleichzeitig überprüft werden. Die Systeme führen dann einen Abgleich durch und nur, wenn beide Systeme zum gleichen Ergebnis kommen und keine Gefährdung erkennen, darf gefahren werden.

”

**Systeme auf der Basis künstlicher Intelligenz bieten zumindest nach derzeitigem Stand keine hundertprozentige Sicherheit. Dennoch können sie wertvolle Ergänzungen darstellen, die zu einer erhöhten Sicherheit führen.“**

Hermann Haase

#### Was ist bei der Entwicklung von KI-Systemen besonders zu beachten?

**STEIMERS:** Die Datenqualität, die in Grundsatz 5 angesprochen wird, ist entscheidend für maschinelles Lernen. Ohne gute Daten kann kein gutes KI-System entstehen. Der Algorithmus erstellt mit den Daten ein

Modell, das bestenfalls alle grundlegenden Eigenschaften einer bestimmten Klasse enthält. Soll ein KI-System lernen, Personen zu erkennen, muss es mit Fotos von Menschen trainiert werden. Alle Fälle, die im echten Leben auftreten können, müssen bedacht werden. Um ein paar Beispiele zu nennen: Menschen mit allen Hautfarben, allen Geschlechtern, Kinder, Menschen im Rollstuhl, bis hin zu verkleideten Menschen. Ein sogenannter Data Scientist hat die Aufgabe, die Datenqualität sicherzustellen und mit der Datenauswahl und -aufbereitung die Realität möglichst nah abzubilden – vollständig ist dies fast nicht möglich.

Die Prüfung und Zertifizierung eines KI-Systems ist zudem nur möglich, wenn es sich um ein deterministisches System handelt: Das Modell des KI-Systems wird zumeist nur in der Entwicklungsphase trainiert und gilt anschließend als fertiggestellt. Nach jeder Anpassung oder Änderung eines KI-Systems muss eine Kontrolle durch den Menschen erfolgen, um die Sicherheit zu gewährleisten. Wenn

das fertige KI-System im Markt etabliert ist, empfiehlt es sich, ähnlich wie in der Medizintechnik oder Pharmazie bei neu eingeführten Medikamenten, weiter zu beobachten und systematisch zu dokumentieren, ob und wann Unfälle geschehen. So kann nachträglich optimiert werden. Modelle, die sich nach der Implementierung automatisch stetig weiterentwickeln und verändern, sogenannte weiterlernende Systeme, können wir nicht prüfen und zertifizieren. Denn wir können nicht wissen, wie sich das Modell nach dem Zeitpunkt der Prüfung und Zertifizierung verhält.

#### Wie kann der Bereich Prüfung und Zertifizierung helfen, die Entwicklung von KI sicher zu gestalten?

**HAASE:** Hier sehe ich insbesondere zwei wichtige Möglichkeiten: Erstens klare Richtlinien, sozusagen „Leitplanken“, zu setzen, mit denen gearbeitet werden kann. Und zweitens durch entwicklungsbegleitende Prüfungen Rückmeldungen zur Sicherheit eines KI-Systems zu geben. Dadurch können Fehlentwicklungen, die die Sicherheit und Gesundheit in unseren

Foto: Privat



Hermann Haase, DGUV Test Projektgruppe Künstliche Intelligenz

”

Wenn das fertige KI-System im Markt etabliert ist, empfiehlt es sich, ähnlich wie in der Medizintechnik oder Pharmazie bei neu eingeführten Medikamenten, weiter zu beobachten und systematisch zu dokumentieren, ob und wann Unfälle geschehen. So kann nachträglich optimiert werden.“

Dr. André Steimers

Mitgliedsbetrieben gefährden, frühzeitig verhindert werden.

**Gibt es inzwischen eine allgemeingültige Definition von KI?**

**STEIMERS:** Eine allgemeingültige Definition von KI ist im Moment noch nicht in

Sicht. Es ist schwierig, alle Methoden, die zum Bereich der KI zählen, in einer Definition zu erfassen und dabei eher triviale Techniken wie einen Taschenrechner auszuschließen. Trotzdem bewegt sich viel auf internationaler Ebene und es lohnt sich, Forschungsprojekte und Entwicklungen zu beobachten. Im Auftrag der DGUV wirke ich in einigen Normungsgremien auch selbst mit – zum Beispiel ISO/IEC JTC 1/SC 42 oder CEN/CLC JTC 21. Aktuell werden viele Standards, beispielsweise zur Datenqualität geschrieben – diese sollten wir kennen und prüfen, was für uns verwendbar ist.

**Beim Thema KI sind noch viele Fragen offen. Ist KI zu komplex für die Praxis oder halten Sie es für eine wichtige Technologie der Zukunft?**

**HAASE:** Es ist wie mit allen Technologien: Am Anfang gibt es viele Fragen, verschiedene Meinungen und natürlich Unzulänglichkeiten. Dies darf aber kein Argument gegen künstliche Intelligenz darstellen. Die ersten Autos, die es gab, sind aus heutiger Sicht absolut unzulänglich gewesen.

Künstliche Intelligenz bietet neue Möglichkeiten und es wird starke Verbesserungen geben, je mehr Anwendungen im Feld sind. In den DGUV Test Prüf- und Zertifizierungsstellen erhalten wir inzwischen häufig Anfragen mit KI-Bezug. Diese Prozesse durch entwicklungsbegleitende Prüfungen voranzubringen und zu unterstützen, stellt eine große Chance dar und ist unsere Aufgabe. Dafür benötigen wir natürlich entsprechende Ressourcen und Expertise in den Prüfstellen.

Das Interview führte Anke Barth, Abteilung Sicherheit und Gesundheit der DGUV (SiGe).

Foto: Sandra Seifen Fotografie



Dr. André Steimers, DGUV Test Projektgruppe Künstliche Intelligenz