Künstliche Intelligenz als Herausforderung für Zulassungsprozesse im Eisenbahnverkehr

Künstliche Intelligenz ermöglicht die prädiktive Instandhaltung und erhöht die Effizienz des Bahnbetriebs. Leistungsfähige Algorithmen zur Umfeldwahrnehmung bilden die Grundlage der Automatisierung des Bahnbetriebs. Die Anwendung nicht-deterministischer Algorithmen und für den Menschen undurchsichtiger Entscheidungsprozesse erfordert einen Paradigmenwechsel in der Zulassung technischer Systeme.

PD Dr.-Ing. habil. Lars Schnieder
Chief Executive Officer (CEO), ESE Engineering und Software-Entwicklung GmbH
lars.schnieder@ese.de

Digitale Technologien dringen immer weiter in alle Bereiche des Eisenbahnsystems vor. Dies schließt auch sicherheitsrelevante Funktionen mit ein. Eingesetzte Technologien müssen vertrauenswürdig sein. Vertrauenswürdige Systeme der künstlichen Intelligenz (nachfolgend als KI-Systeme bezeichnet) sind Grundlage für eine höhere Wettbewerbsfähigkeit des Verkehrsträgers Schiene. Ein KI-System ist eine Software, „die mit einer oder mehreren Techniken und Konzepten entwickelt worden ist und im Hinblick auf eine Reihe von Zielen, die vom Menschen festgelegt werden, Ergebnisse wie Inhalte, Vorhersagen, Empfehlungen oder Entscheidungen hervorbringen kann, die das Umfeld beeinflussen, mit dem sie interagieren". Bei der Einführung von KI-Systemen stellen sich produkthaftungsrechtliche Fragen. Im Falle KI-gestützter Systeme ist es jedoch schwer, einen Produktfehler, den entstandenen Schaden und den Kausalzusammenhang nachzuweisen [1]. Dieser Beitrag skizziert Schutzziele, produkthaftungsrechtliche Fragestellungen, sowie den von der Kommission der Europäischen Union skizzierten zukünftigen Regelungsrahmen für KI-Systeme.

1. Schutzziele bei Einsatz von Algorithmen der Künstlichen Intelligenz

Die Potenziale von Algorithmen der Künstlichen Intelligenz sind weitreichend. Allerdings resultieren aus der Einführung dieser neuen Technologien auch Risiken. Diese müssen durch angemessene Maßnahmen des Regelungsrahmens beherrscht werden. Konkret umfasst dies die folgenden Schutzziele:

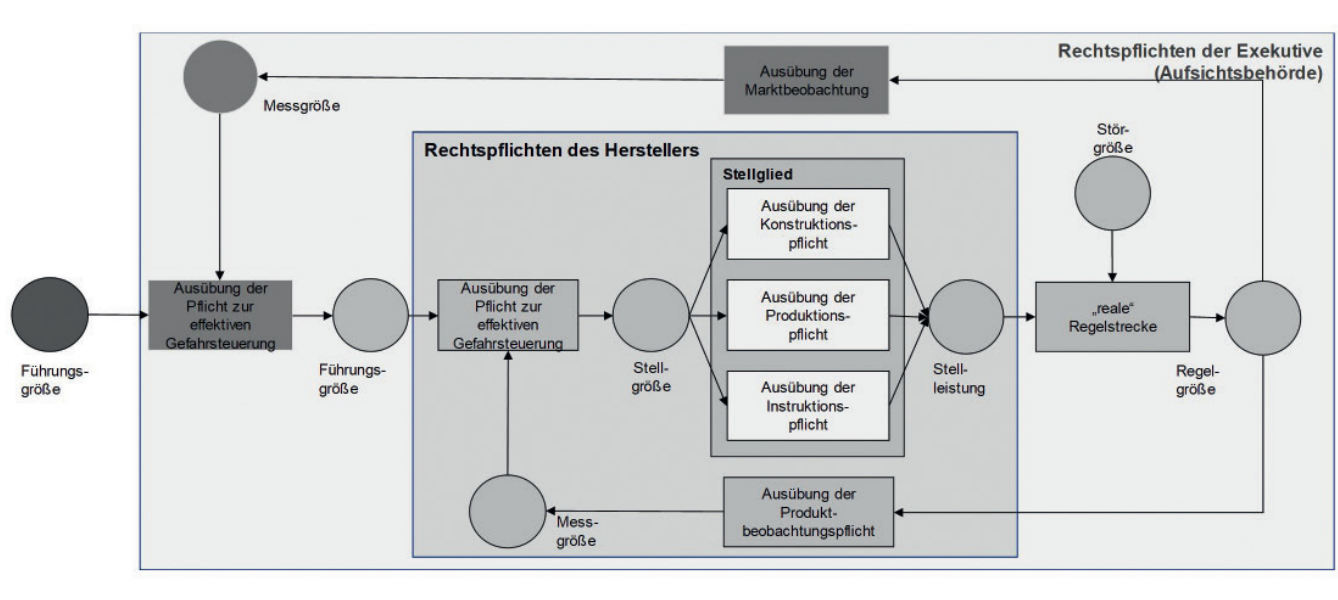
- *Wahrung der Grundrechte:* Für den Einsatz von Algorithmen der Künstlichen Intelligenz bestehen hohe Hürden hinsichtlich ethischer Belange. Der Einsatz von Algorithmen der Künstlichen Intelligenz darf zu keinen nachteiligen Auswirkungen auf die Grundrechte natürlicher Personen führen. Beispiele hierfür sind der Schutz der Privatsphäre, der Schutz personenbezogener Daten, aber auch der Grundsatz der Diskriminierungsfreiheit.
- *Sicherheit:* Technische Systeme müssen frei von nicht akzeptierten Risiken sein. Dieser Anspruch an die Integrität technischer Systeme erfordert bereits bei der Entwicklung konventioneller Systeme umfassende Maßnahmen zur Vermeidung systematischer Fehler und zufälliger Ausfälle. Kommen KI-Systeme mit für den Menschen undurchsichtigen Entscheidungsprozessen zum Einsatz, erschwert dies den Nachweis ihres bestimmungsgemäßen Verhaltens [2].
- *Rechtssicherheit:* Rechtssicherheit beruht auf dem Anspruch der Klarheit, Beständigkeit, Vorhersehbarkeit und Gewährleistung von Rechtsnormen sowie die an diese gebundenen konkreten Rechte und Pflichten. Unternehmen investieren nur dann in technologische Innovationen, wenn die hieraus resultierenden (rechtlichen) Risiken für sie beherrschbar sind.

2. Rechtspflichten der Hersteller

Im Binnenmarkt der Europäischen Union bestehen hohe Hürden für das Inverkehrbringen technischer Systeme – insbesondere Bahnanwendungen. In der Entwicklung sicherer Bahnanwendungen müssen von den Herstellern umfassende Rechtspflichten beachtet werden. Die Rechtspflichten der Hersteller [3] bei Einführung von KI-Systemen bilden eine geschlossene und rückgekoppelte Wirkstruktur (Bild 1). Dieser Kreislauf ist in einen übergeordneten Regelkreis staatlicher Marktbeobachtung eingebettet. Dieser aus der europäischen und nationalen Rechtssetzung und Rechtsprechung gesetzte Rahmen (vgl. [4]) wird nachfolgend umrissen.

2.1. Pflicht zur effektiven Gefahrsteuerung

Dreh- und Angelpunkt der Produktentwicklung mit Methoden der KI ist die *Pflicht zur effektiven Gefahrsteuerung* (Regelein-



1: Rechtspflichten der Hersteller als rückgekoppelte Wirkstruktur [5]

richtung). Hierbei werden Messgrößen gegen die vorgegebenen Führungsgrößen (angestrebtes Sicherheitsniveau im Eisenbahnsystem) verglichen und gemeldete Vorfälle hinsichtlich ihrer Sicherheitsrelevanz bewertet. Der Hersteller muss ein Risikomanagementsystem einrichten, anwenden, dokumentieren und aufrechterhalten. Das Risikomanagementsystem ist ein kontinuierlicher iterativer Prozess während des gesamten Lebenszyklus eines KI-Systems. In der Entwurfsphase sind bekannte und vorhersehbare Risiken zu ermitteln [6]. Basierend hierauf sind bei der gemäß seiner Zweckbestimmung oder im Rahmen einer vernünftigerweise vorhersehbaren Fehlanwendung entstehende Risiken abzuschätzen und zu bewerten. Mögliche nach Inverkehrbringen im Zuge der Marktbeobachtung entstehende Risiken sind zu bewerten [7] und hierfür bei Bedarf Maßnahmen zur Verringerung des Risikos zu ergreifen [8].

2.2. Konstruktionspflicht

In der Entwicklung des Produkts (Stellglied) bestehen verschiedene Möglichkeiten zur gezielten sicherheitsgerichteten Ausgestaltung des KI-Systems. Die *Konstruktionspflicht* fordert vom Hersteller alle im vernünftigen Ermessen stehenden Maßnahmen zur Absicherung des KI-Systems. In der Rechtsprechung des Bundesgerichtshofs [9] hat sich hierbei als Maßstab einer rechtssicheren Zulassung risikobehafteter Technologien der Stand von Wissenschaft und Technik etabliert (vgl. Definition in [10]). Es muss also sichergestellt werden, dass das KI-System die „konstruktiven" Anforderungen des Richtlinienentwurfs [8] erfüllt. Dies umfasst die folgenden Aspekte:

- *Erfüllung von Anforderungen an die Datenqualität:* Eine hohe Datenqualität ist für die Leistung von KI-Systemen wesentlich. Insbesondere wenn Techniken eingesetzt werden, bei denen Modelle mit Daten trainiert werden, um sicherzustellen, dass das KI-System bestimmungsgemäß und sicher funktioniert. Die Trainings-, Validierungs- und Testdatensätze müssen hinsichtlich der Zweckbestimmung des Systems relevant, repräsentativ, fehlerfrei und vollständig sein [8].
- *Erfüllung von Anforderungen an die menschliche Aufsicht:* KI-Systeme sollten so konzipiert und entwickelt werden, dass natürliche Personen ihre Funktionsweise überwachen können. Die natürlichen Personen, denen die menschliche Aufsicht übertragen wird, müssen über die erforderliche Kompetenz, Ausbildung und Befugnis verfügen [8].
- *Erfüllung von Anforderungen an die Genauigkeit, Robustheit und Cybersicherheit:* Erreichte Kennzahlen zur Genauigkeit müssen den Nutzern in der Gebrauchsanweisung mitgeteilt werden [8]. Robustheit bezeichnet eine Widerstandsfähigkeit im Zusammenhang mit den Grenzen des Systems insbesondere in Bezug auf Fehler, Störungen, Unstimmigkeiten sowie unerwartete Situationen. Cybersicherheit erfordert die umfassende Betrachtung möglicher böswilliger Eingriffe. Diese böswilligen Eingriffe dürfen die Integrität des KI-Systems nicht gefährden und zu schädlichem oder anderweitig unerwünschtem Verhalten führen [8]. Dies erfordert Maßnahmen zur Verhütung und Kontrolle von Angriffen, mit denen über Schwachstellen versucht wird, Trainingsdatensätze zu manipulieren (z. B. Datenvergiftung), das KI-System durch Beaufschlagung mit manipulierten Eingabedaten zu Fehlern zu verleiten, oder Modellmängel einzuführen. Außerdem sind Schwachstellen in den digitalen Ressourcen des KI-Systems oder der IKT-Infrastruktur (Informations- und Kommunikationstechnik) zu vermeiden [8].

2.3. Produktionspflicht

Die *Produktionspflicht* fordert vom Hersteller eine umfassende Qualitätsplanung und Qualitätssicherung der angewendeten Fertigungsprozesse. Die Hersteller müssen also ein solides Qualitätsmanagementsystem einrichten und die Durchführung des vorgeschriebenen Konformitätsbewertungsverfahrens auf der Grundlage der

Bewertung des Qualitätsmanagementsystems. Auf diese Weise soll verhindert werden, dass Produkte mit möglicherweise sicherheitskritischen Toleranzen aus dem Fertigungsprozess ins Feld kommen, wo diese Gefährdungen hervorrufen könnten.

2.4. Instruktionspflicht

Darüber hinaus greift die *Instruktionspflicht*. Die Nutzerdokumentation beschreibt den bestimmungsgemäßen Gebrauch des KI-Systems. Zuwiderhandlungen gegen die Vorgaben der Nutzerdokumentation begründen den Fehlgebrauch. Dies ist faktisch eine Haftungsbegrenzung der Hersteller. Dies ist für die Entwicklung von KI-Systemen in zweierlei Hinsicht relevant:

- *Technische Dokumentation:* Informationen darüber, wie KI-Systeme entwickelt wurden und wie diese während ihres gesamten Lebenszyklus funktionieren, sind unerlässlich, um die Einhaltung der Schutzziele überprüfen zu können. Diese Informationen sollten allgemeine Merkmale, Fähigkeiten und Grenzen des Systems, die verwendeten Algorithmen, Daten, Trainings-, Test- und Validierungsverfahren sowie die Dokumentation des einschlägigen Risikomanagementsystems umfassen. Die technische Dokumentation sollte über den Lebenszyklus [11] des KI-Systems stets auf dem neuesten Stand gehalten werden [8].
- *Transparenz und Bereitstellung von Informationen für die Nutzer:* Die Nutzer sollten in der Lage sein, die Ergebnisse des Systems zu interpretieren und es angemessen zu verwenden. KI-Systemen sollte daher einschlägige Dokumentation und Gebrauchsanweisungen beigefügt sein [8]. Die Gebrauchsanweisung muss präzise, vollständige, korrekte und eindeutige Informationen in einer für den Nutzer relevanten, barrierefrei zugänglichen und verständlichen Form enthalten. Insbesondere sind Angaben zu erforderlichen Wartungs- und Pflegemaßnahmen zur Gewährleistung eines ordnungsgemäßen Funktionierens, auch in Bezug auf Software-Updates, erforderlich.

2.5. Produktbeobachtungspflicht

Es schließt sich das Inverkehrbringen des KI-Systems an. Von diesem Zeitpunkt an zeigt sich die (Sicherheits-)Leistungsfähigkeit des Produktes am Markt (Regelstrecke). Das KI-gestützte System wird hierbei vom Eisenbahnunternehmen in der Regel bestimmungsgemäß betrieben. Dennoch können im Betriebsablauf Unregelmäßigkeiten auftreten. Spätestens mit dem im Produkthaftungsrecht wegweisenden Honda-Urteil des Bundesgerichtshofes [12] haben die Hersteller umfassende Verfahren und Verantwortlichkeiten in der Ausübung ihrer Produktbeobachtungspflicht klar geregelt. Dies spiegelt sich auch im Richtlinienentwurf [8]. Demnach sind weitreichende Aufzeichnungspflichten einzuhalten. KI-Systeme werden mit Funktionsmerkmalen konzipiert und entwickelt, die eine automatische Aufzeichnung von Vorgängen und Ereignissen („Protokollierung") während des Betriebs der KI-Systeme ermöglichen. Die Protokollierung ermöglicht die Überwachung des Betriebs des KI-Systems im Hinblick auf das Auftreten von Situationen, die dazu führen können, dass das KI-System ein Risiko birgt. Dies erleichtert die Beobachtung nach dem Inverkehrbringen.

3. Europäischer Regulierungsrahmen für KI-Systeme

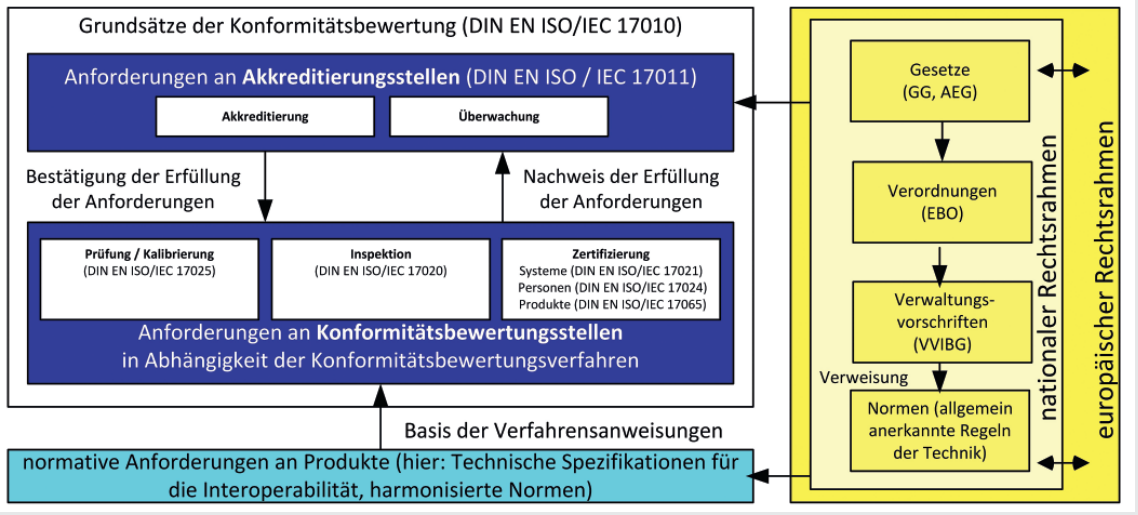
Ein klarer europäischer Regulierungsrahmen stärkt das Vertrauen in die künstliche Intelligenz und beschleunigt damit die Einführung der Technologie [1]. Was genau sind die Elemente des Regulierungsrahmens?

- *Harmonisierte Normen:* Normen sind von einer anerkannten Normungsorganisation angenommene technische Spezifikationen zur wiederholten oder ständigen Anwendung. Ihre Einhaltung ist im Rechtsverkehr nicht zwingend, aber aus rechtlichen Gründen (bspw. Beweislastumkehr) geboten. Harmonisierte Normen sind europäische Norm, die auf der Grundlage eines Auftrags der Kommission zur Durchführung von Harmonisierungsrechtsvorschriften der Union angenommen wurden [13].
- *Konformitätsbewertung:* KI-Systeme sind im Rahmen der bestehenden Konformitätsbewertungsverfahren zu betrachten. Hierbei handelt es sich um die Prüfung, Inspektion oder Zertifizierung (vgl. Bild 2). Die Konformitätsbewertung umfasst hierbei sowohl die Algorithmen als auch die in der Entwicklungsphase verwendeten Datensätze. Wichtig ist die Durchführung einer erneuten Konformitätsbewertung bei wesentlichen Änderungen der KI-Systeme. Es gelten auch in Bezug auf KI-Systeme die an Konformitätsbewertungsstellen gerichteten Anforderungen zur Unabhängigkeit, Objektivität und Unparteilichkeit [14].
- *Marktüberwachung:* Es sind über die Ex-ante zu berücksichtigenden konstruktiven Merkmale hinaus auch Ex-Post-Mechanismen für die Rechtsbefolgung und -durchsetzung erforderlich. Dies bezeichnet insbesondere Verfahren zur Meldung schwerwiegender Vorfälle und Fehlfunktionen. Hierbei unterliegen die Hersteller von KI-Systemen bei schwerwiegenden Vorfällen oder Fehlfunktionen von KI-Systemen einer Meldepflicht. Diese Meldung erfolgt unmittelbar, nachdem der Anbieter den kausalen Zusammenhang zwischen dem KI-System und dem Vorfall bzw. der Fehlfunktion oder die naheliegende Wahrscheinlichkeit eines solchen Zusammenhangs festgestellt hat.

4. Fazit und Ausblick

Die Kommission der Europäischen Union hat den Regelungsbedarf in Bezug auf die rechtssichere Einführung von KI-Systemen erkannt und hierfür einen Richtlinienentwurf vorgelegt. Dieser Richtlinienentwurf ist für die Rechtssicherheit notwendig, aber nicht hinreichend. Zwar wird die Herstellerverantwortung hinsichtlich der einzelnen von diesen wahrzunehmenden Rechtspflichten näher konkretisiert und der allgemein gültige Regelungsrahmen für die Einführung risikobehafteter Technologien im Grundsatz bestätigt, jedoch fehlt mit einer harmonisierten Norm das Rückgrat für die wirksame Umsetzung des von der Kommission skizzierten Konzepts. In Bezug auf KI-Systeme stößt die Normung an die folgenden Grenzen:

- **Normen sind nicht allein entscheidend:** Normen enthalten im Allgemeinen keine abschließenden Verhaltensanforderungen und bestimmen nicht die Grenze dessen, was im Einzelfall verlangt werden kann. Es ist daher grundsätzlich von den Herstellern selbständig zu prüfen, ob und welche Sicherungsmaßnahmen im Zusammenhang mit KI-Systemen notwendig sind [15].
- **Normen können unzutreffend sein (Widerlegung der Richtigkeitsvermutung):** Es kann (muss aber nicht) davon ausgegangen werden, dass die Norm das wiedergibt, was der Gesetzgeber fordert (sog. Richtigkeitsvermutung). Die Rich-



2: Systematik der Konformitätsbewertungsstellen [14]

tigkeitsvermutung ist aber widerlegbar, da Normen „falsch" oder „unzutreffend" sein können. Da die aktuelle Normungslandschaft für die Software für Bahnanwendungen [16] auf deterministischen Algorithmen fokussiert, besteht hier Handlungsbedarf.

- **Normen können unvollständig sein (keine Vollständigkeitsvermutung):** Normen regeln die Verkehrssicherungspflichten nicht abschließend. Die Einhaltung von Normen ist daher notwendig, aber nicht hinreichend. Es muss beispielsweise auch der vorhersehbare Fehlgebrauch in der Konstruktion von KI-Systemen mitberücksichtigt werden.
- **Normen können veraltet sein (keine Aktualitätsvermutung):** Normen bieten einen Anhaltspunkt für den Stand der Technik zum Zeitpunkt der Bekanntmachung. Normen gelten aus rechtlicher Sicht nicht so lange, bis sie zurückgezogen werden, sondern nur solange sie das Gesetz noch zutreffend konkretisieren. Diese Grenze ist insofern relevant, als dass sich zum Zeitpunkt der Veröffentlichung der Norm für Software für Bahnanwendungen KI-Systeme noch nicht abzeichneten.

Der erste Schritt eines Regelungsrahmens für KI-Systeme ist getan. Es liegt jetzt an der Bahnbranche, die bestehenden Unschärfen durch die aktive Mitwirkung in der Normung mit konkreten Vorgaben zu schließen. Dann besteht eine realistische Chance, im nächsten Jahrzehnt den Verkehrsträger Schiene im intermodalen Wettbewerb zu stärken. ●

Literatur

[1] Kommission der Europäischen Union: Weissbuch „Zur Künstlichen Intelligenz – ein europäisches Konzept für Exzellenz und Vertrauen"; COM(2020) 65 final; 19.02.2020.
[2] DIN EN 50129:2019-06: Bahnanwendungen; Telekommunikationstechnik, Signaltechnik und Datenverarbeitungssysteme – Sicherheitsbezogene elektronische Systeme für Signaltechnik; Deutsche Fassung EN 50129:2018 + AC:2019
[3] Klindt, T.; Handorn, B. (2010): Haftung eines Herstellers für Konstruktions- und Instruktionsfehler. Neue Juristische Wochenschrift Jahrgang 63 (2010) Ausgabe 16, S. 1105 – 1108
[4] Lukes, Rudolf (Hrsg.): Sicherheitsverantwortung im Eisenbahnwesen. Carl Heymanns Verlag (Köln) 2002.
[5] Schnieder, Lars; Hosse, René: Typgenehmigungsmaßstäbe für automatisierte Fahrzeugsysteme des Level 3. in: Zeitschrift für Verkehrssicherheit (*) 65 (2019) 4, S. 246 – 252.
[6] DIN EN 50126-2:2018-10: Bahnanwendungen – Spezifikation und Nachweis von Zuverlässigkeit, Verfügbarkeit, Instandhaltbarkeit und Sicherheit (RAMS) – Teil 2: Systembezogene Sicherheitsmethodik; Deutsche Fassung EN 50126-2:2017
[7] DIN VDE V 0831-100:2019-08: Elektrische Bahn-Signalanlagen Teil 100: Risikoorientierte Beurteilung von potenziellen Sicherheitsmängeln und risikoreduzierende Maßnahmen
[8] Kommission der Europäischen Union: Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Gesetz über künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union. COM(2021) 206 final, 21.04.2021.
[9] Bundesgerichtshof. Urteil vom 16.6.2009, Aktenzeichen VI ZR 107/08. (Haftung eines Fahrzeugherstellers für die Fehlauslösung von Airbags)
[10] Bundesministerium der Justiz. 2008. Handbuch der Rechtsförmlichkeit. Berlin: BMVJ
[11] DIN EN 50126-1:2018-10: Bahnanwendungen – Spezifikation und Nachweis von Zuverlässigkeit, Verfügbarkeit, Instandhaltbarkeit und Sicherheit (RAMS) – Teil 1: Generischer RAMS-Prozess; Deutsche Fassung EN 50126-1:2018
[12] Bundesgerichtshof. Urteil vom 9.12.1986, Aktenzeichen VI ZR 65/86 (Motorrad-Lenkerverkleidung)
[13] Verordnung (EU) Nr. 1025/2012 des Europäischen Parlaments und des Rates vom 25. Oktober 2012 zur europäischen Normung, zur Änderung der Richtlinien 89/686/EWG und 93/15/EWG des Rates sowie der Richtlinien 94/9/EG, 94/25/EG, 95/16/EG, 97/23/EG, 98/34/EG, 2004/22/EG, 2007/23/EG, 2009/23/EG und 2009/105/EG des Europäischen Parlaments und des Rates und zur Aufhebung des Beschlusses 87/95/EWG des Rates und des Beschlusses Nr. 1673/2006/EG des Europäischen Parlaments und des Rates
[14] Schnieder, Lars: Öffentliche Kontrolle der Qualitätssicherungskette für einen sicheren und interoperablen Schienenverkehr. in: ETR-Eisenbahntechnische Rundschau 66 (2017) 4, S. 38 – 41.
[15] Wilrich, Thomas: Die rechtliche Bedeutung technischer Normen als Sicherheitsmaßstab. 1. Auflage. Beuth Verlag (Berlin) 2017
[16] DIN EN 50128:2012-03: Bahnanwendungen – Telekommunikationstechnik, Signaltechnik und Datenverarbeitungssysteme – Software für Eisenbahnsteuerungs- und Überwachungssysteme; Deutsche Fassung EN 50128:2011

Summary

Artificial Intelligence as challenge for approval processes in the railway transport

AI allows predictive maintenance and increases the efficiency of the rail system. Efficient algorithms for environment perception are the basis for automating the railway operation. The application of non-deterministic algorithms and decision-making processes that are opaque to humans require a paradigm shift in the approval process of technological systems.