EnBW Rückmeldung zu EU-KOM

Gesetzesvorschlag zu Künstlicher Intelligenz

Kontext

Aus Sicht der EnBW Energie Baden-Württemberg AG wurde mit dem Gesetzesvorschlag der EU-Kommission zu Künstlicher Intelligenz (KI), eine Regulierung entworfen, welche Chancen eröffnet. Nach erster Betrachtung, hiermit unsere erste kurze Rückmeldung aus energiewirtschaftlicher Perspektive sowie dem Betrieb von kritischer Infrastruktur.

Hauptempfehlungen

Während der in weiten Teilen innovationsoffene Charakter des VO-Vorschlags sehr zu begrüßen ist, könnten nachfolgende Punkte Klarheit und Ziele der VO weiter unterstützen.

- Definition von „safety components" (Annex III; 2.a): Eine Abgrenzung von Sicherheitskomponenten könnte weitergehend Klarheit erzeugen. Bsp. ob Netzplanung, Personal- und Schichtplanung im weitesten Sinne als Sicherheitskomponente zählt. Dass KI in der Anlagensteuerung im Kraftwerk oder der Netzleitstelle unter die Kategorie „high risk" fällt, ist nachvollziehbar – eine Abgrenzung oder Abstufung könnte hier weiter präzisieren. [→ siehe Abbildung mit Anwendungsfällen auf Seite 4]

- Beispielhafte technische Dokumentation zur Präzision der Anforderungen (Annex IV): Da die Erwartung hinsichtlich des Detailgrades relevant, aber nicht klar ist, wäre an dieser Stelle die Zurverfügungstellung eines konkreten Beispiels sinnvoll.

- Anlaufstelle für Fragen zu KI-Anwendungen: Für Rückfragen zu konkreten KI-Systemen in Entwicklung oder im Ideenstatus wäre es hilfreich (nationale) Anlaufstelle(n) zu haben, mit der man Fragen diskutieren kann. Besonders die Einordnung ob eine KI-Anwendung „high-risk" oder nicht ist, wäre ist hier wichtig.

- KI-Systeme zur internen Anwendung: Aus Sicht des Anlagenbetriebs (bspw. in der Energiewirtschaft) werden KI-Systeme absehbar, im Wesentlichen, intern verwendet und nicht als „Produkt auf einen Markt und in Umlauf gebracht". Großteils/Weniger (Produkt-)Endkunden, sondern Mitarbeiter sind Nutzer der Systeme (Bsp. KI-basierte Instandhaltungsplanung von Gasnetzen). Eine Abgrenzung von Anwendungsbereichen von KI könnte hier sinnvoll sein.

Detaillierte Einschätzung

1. Aus Sicht EnBW positiv und zu begrüßen:

Die Möglichkeit zu Inhouse conformity assessments werten wir als innovationsfördernd. Aufwand und Hemmnisse werden hierdurch wesentlich reduziert, da keine Abstimmung mit Dritten notwendig ist. Durch interne Prüfungen findet im Unternehmen ebenfalls Wissens- und Kompetenzaufbau bzgl. der Konformität statt.

- Seite 33: (64) *„Therefore, the conformity assessment of such systems should be carried out as a general rule by the provider under its own responsibility, with the only exception of AI systems intended to be used for the remote biometric identification of persons, for which the involvement of a notified body in the conformity assessment should be foreseen, to the extent they are not prohibited."*

2. Aus Sicht EnBW besteht Klärungsbedarf:

1. Bestimmung von Hochrisiko-Anwendungen;

- Seite 26: (34) und Annex III, 2a *„As regards the management and operation of critical infrastructure, it is appropriate to classify as high-risk the AI systems intended to be used as safety components in the management and operation of road traffic and the supply of water, gas, heating and electricity, since their failure or malfunctioning may put at risk the life and health of persons at large scale and lead to appreciable disruptions in the ordinary conduct of social and economic activities."*

Für uns wäre eine Hoch-Risiko Anwendung durch KI, bspw. eine autonome (Strom-)Netzsteuerung. Keine Hoch-Risikoanwendungen wären bspw. KI-basierte Erkennung von Schäden an Masten und Freileitungen oder Handelsalgorithmen (z.B. im physischen kurzfristigen Stromhandel).

- Seite 18: *"The definition should be based on the key functional characteristics of the software, in particular the ability [...] to generate [...] which influence the environment with which the system interacts"*

In diesem Kontext müsste "influence" bzw. beeinflussen konkretisiert werden. Wenn ein System Controller eine (auf statistischen Methoden basierte) Trafolast für seine Schaltvorgänge verwendet, wäre dies dann bereits KI/AI?

- Seite 24: (27) *„AI systems identified as high-risk should be limited to those that have a significant harmful impact on the health, safety and fundamental rights of persons in the Union"*

Hier wäre eine genauere Definition hilfreich. Wie direkt muss der Impact sein? Welche Mittelbarkeit/Unmittelbarkeit muss vorliegen? Beispiel: In der KI-basiertes Instandhaltungsplanung, trifft die KI eine falsche Entscheidung. Ein deshalb nicht ausgetauschtes Bauteil sorgt für längere Zeit für einen Stromausfall. Die

Wahrscheinlichkeit ist sehr gering, dass eine Verkettung von Ereignissen vorliegt. Es ergibt sich die Frage, ob solche Konsequenzen berücksichtigt werden sollten.

2. Teilweise recht offene bzw. weite Formulierungen in der KI-VO wie „a certain degree", „appropriate" würden wir als unternehmerische Freiheit und Verantwortung bei der Adressierung der Anforderungen auffassen.

- Seite 30: (47) *„To address the opacity that may make certain AI systems incomprehensible to or too complex for natural persons, a certain degree of transparency should be required for high-risk AI systems."*

- Seite 29 (44): " *Training, validation and testing data sets should be sufficiently relevant, representative and free of errors and complete in view of the intended purpose of the system."*

3. Hinsichtlich der Technical Documentation wäre eine beispielhafte Dokumentation sehr hilfreich.

- Annex IV, Technical Documentation: *„The provider verifies that the established quality management system is in compliance with the requirements of Article 17"*

Da die Erwartung hinsichtlich des Detailgrades relevant, aber nicht klar ist, wäre an dieser Stelle die Zurverfügungstellung eines konkreten Beispiels hilfreich.

4. Intern genutzte KI-Systeme vs. externe in Verkehr gebrachte Produkte. Besonders aus der Perspektive der Energiewirtschaft und kritischer Infrastruktur werden KI-Systeme und Lösungen oft im Wesentlichen, intern verwendet und nicht als Produkt auf einen Markt und in Umlauf gebracht. Ein Beispiel ist die KI-basierte Instandhaltungsplanung von Gasnetzen. In diesem und vielen weiteren Fällen sind Mitarbeiterinnen und Mitarbeiter die Nutzerinnen und Nutzer des KI-Systems. Uns ist noch nicht klar, inwieweit die Regulierung hier eine Unterscheidung von Endkundenprodukten vorsieht. Wie genau funktioniert in den Fällen der rein internen Nutzung eine Market Surveillance?

5. Die Definition von KI in Annex 1 ist sehr breit angelegt. Die Definition (incl. Expert Systems und Statistical Approaches) würde in dieser Form bereits existierende mathematische Lösungen und deterministische Ansätze (z.B. stochastische Prozesse), unter die Definition von KI fallen lassen. Diese erfordern aber nicht die gleichen Qualitätssicherungsmaßnahmen wie lernende Algorithmen.

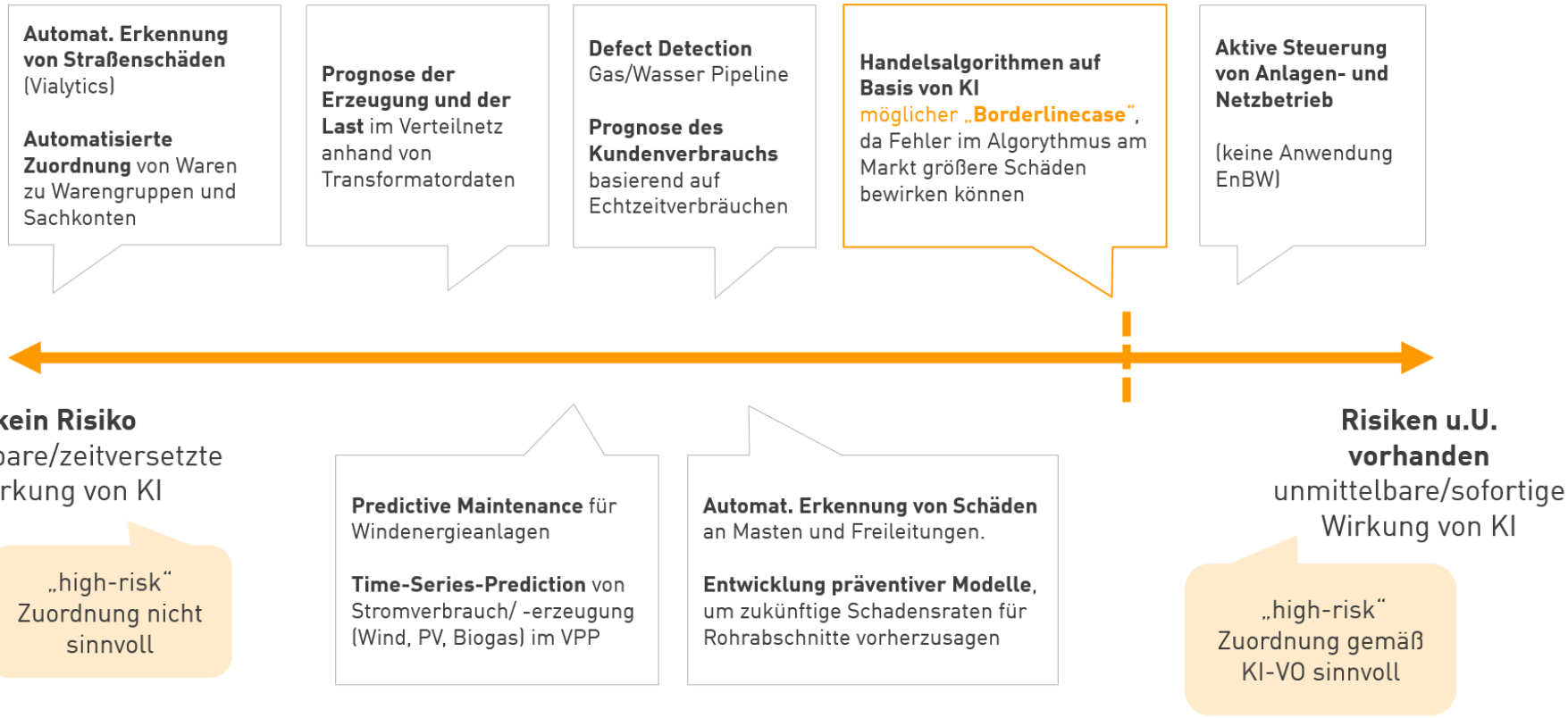
3. Anregungen aus Sicht EnBW:

Anlaufstelle bei Unklarheiten: Generell bleibt es dem Entwickler/Betreiber selbst überlassen zu interpretieren, ob ein AI-System high-risk ist. Es wäre hilfreich eine Anlaufstelle zu haben, mit der man diese Frage diskutieren kann. Alternativ müsste man jeglichen Interpretationsspielraum entfernen, was nicht sinnvoll möglich ist.

Höhe der Sanktionen (Article 71): Bei gefährdendem Verhalten, ganz besonders im Falle des Einsatzes von KI der„unacceptable risk"-Klasse, kann ein hohes Strafmaß gerechtfertigt sein. Hohe Strafmaße von bis zu 6% des weltweiten Unternehmensumsatzes (bis zu 30 Mio €) hingegen, können wohlmöglich grundlegend

abschreckend wirken. Hohe Sanktionen können problematisch sein, wenn Fehler z.B. auf Grund von Unklarheiten der Regulierung geschehen und schwer geahndet werden. Bei bewussten Täuschungen und gefährdendem Verhalten können die Sanktionen jedoch angemessen und wichtig sein.

Risikospektrum von KI-Systemen in Energiewirtschaft & kritischer Infrastruktur (Use-Cases EnBW)



2