

# Die Definition eines KI-Systems – Drei-Faktor-Ansatz

Univ.-Prof. Dr. Christiane Wendehorst, LL.M. (Cambridge)



- 1 Entwicklung der Definition von „KI-System“
- 2 Elemente der Definition in Art. 3 Nr. 1 KI-VO
- 3 Drei-Faktor-Ansatz

1

# Entwicklung der Definition eines KI-Systems

## HLEG 2019



### GLOSSAR

(142) Dieses Glossar ist Bestandteil dieser Leitlinien und dient dem Verständnis der in diesem Papier verwendeten Begriffe.

#### **Künstliche Intelligenz oder KI-Systeme**

(143) Künstliche-Intelligenz-(KI)-Systeme sind vom Menschen entwickelte Software- (und möglicherweise auch Hardware-) Systeme<sup>78</sup>, die in Bezug auf ein komplexes Ziel auf physischer oder digitaler Ebene agieren, indem sie ihre Umgebung durch Datenerfassung wahrnehmen, die gesammelten strukturierten oder unstrukturierten Daten interpretieren, Schlussfolgerungen daraus ziehen oder die aus diesen Daten abgeleiteten Informationen verarbeiten und über die geeignete(n) Maßnahme(n) zur Erreichung des vorgegebenen Ziels entscheiden. KI-Systeme können entweder symbolische Regeln verwenden oder ein numerisches Modell erlernen, und sie können auch ihr Verhalten anpassen, indem sie analysieren, wie die Umgebung von ihren vorherigen Aktionen beeinflusst wird.

(144) Als wissenschaftliche Disziplin umfasst die KI mehrere Ansätze und Techniken wie z. B. maschinelles Lernen (Beispiele dafür sind „Deep Learning“ und bestärkendes Lernen), maschinelles Denken (es umfasst Planung, Terminierung, Wissensrepräsentation und Schlussfolgerung, Suche und Optimierung) und die Robotik (sie umfasst Steuerung, Wahrnehmung, Sensoren und Aktoren sowie die Einbeziehung aller anderen Techniken in cyber-physische Systeme).

(145) In einem von der HEG-KI separat erarbeiteten und parallel zu dem vorliegenden Dokument veröffentlichten Papier mit dem Titel „Eine Definition der KI: Wichtigste Fähigkeiten und Wissenschaftsgebiete“ wird die Definition des Begriffs *KI-Systeme* für die Zwecke dieser Leitlinien eingehender erläutert.

## EU-Kommission 2021



### Artikel 3 Begriffsbestimmungen

Für die Zwecke dieser Verordnung bezeichnet der Ausdruck

1. „System der künstlichen Intelligenz“ (KI-System) eine Software, die mit einer oder mehreren der in Anhang I aufgeführten Techniken und Konzepte 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;

### **ANHANG I** **TECHNIKEN UND KONZEPTE DER KÜNSTLICHEN INTELLIZENZ** **gemäß Artikel 3 Absatz 1**

- a) Konzepte des maschinellen Lernens, mit beaufsichtigtem, unbeaufsichtigtem und bestärkendem Lernen unter Verwendung einer breiten Palette von Methoden, einschließlich des tiefen Lernens (*Deep Learning*);
- b) Logik- und wissensgestützte Konzepte, einschließlich Wissensrepräsentation, induktiver (logischer) Programmierung, Wissensgrundlagen, Inferenz- und Deduktionsmaschinen, (symbolischer) Schlussfolgerungs- und Expertensysteme;
- c) Statistische Ansätze, Bayessche Schätz-, Such- und Optimierungsmethoden.

## EU-Parlament 2023



---

---

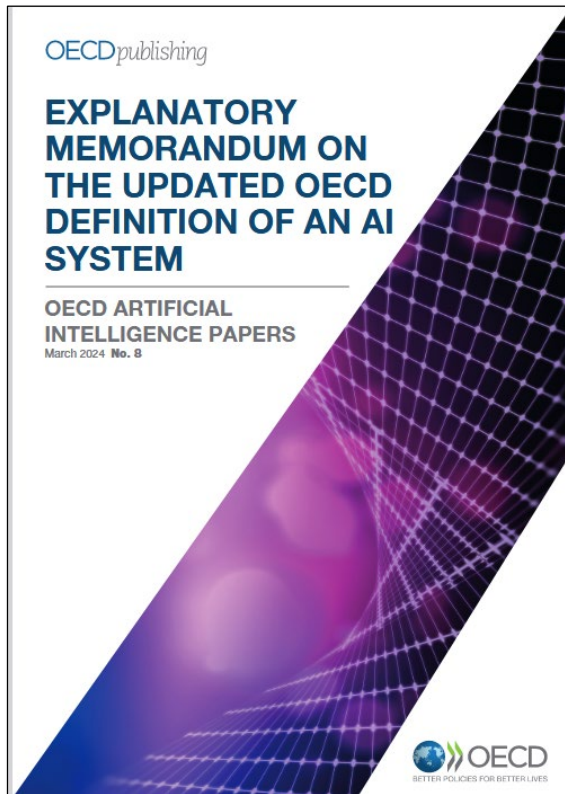
### *Geänderter Text*

---

1. „System der künstlichen Intelligenz“ (KI-System) *ein maschinengestütztes System, das so konzipiert ist, dass es mit unterschiedlichem Grad an Autonomie operieren kann und das für explizite oder implizite Ziele Ergebnisse wie Vorhersagen, Empfehlungen oder Entscheidungen hervorbringen kann, die das *physische oder virtuelle* Umfeld beeinflussen;*

---

## OECD 2019/2023



The OECD definition of an AI system contained in the OECD AI Principles (OECD, 2019<sup>[1]</sup>); (OECD, 2019<sup>[2]</sup>) built on the conceptual view of AI detailed in Artificial Intelligence: A Modern Approach (Russell and Norvig, 2009<sup>[3]</sup>). It read: *“An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy”*.

The text above is replaced with the following updated definition:

*An AI system is a machine-based system that can, for a given set of human-defined explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as makes predictions, content, recommendations, or decisions that can influence physical real or virtual environments. Different AI systems are designed to operate with varying in their levels of autonomy and adaptiveness after deployment.*

The updated definition reads as follows:

***An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.***

# Endgültige Fassung Art. 3 Nr. 1 KI-VO

## *Artikel 3*

### **Begriffsbestimmungen**

Für die Zwecke dieser Verordnung bezeichnet der Ausdruck

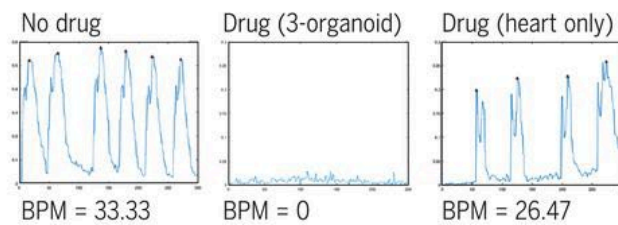
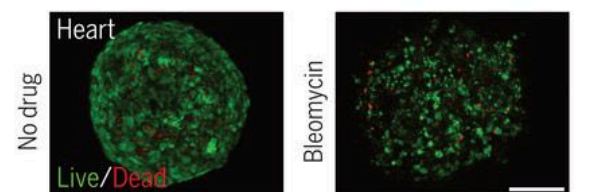
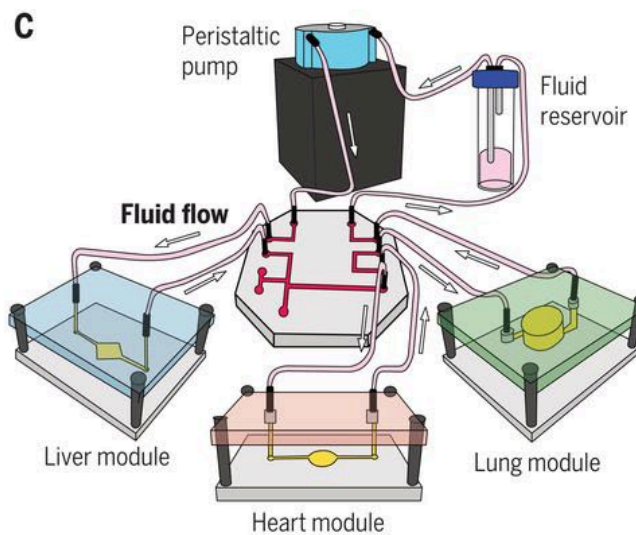
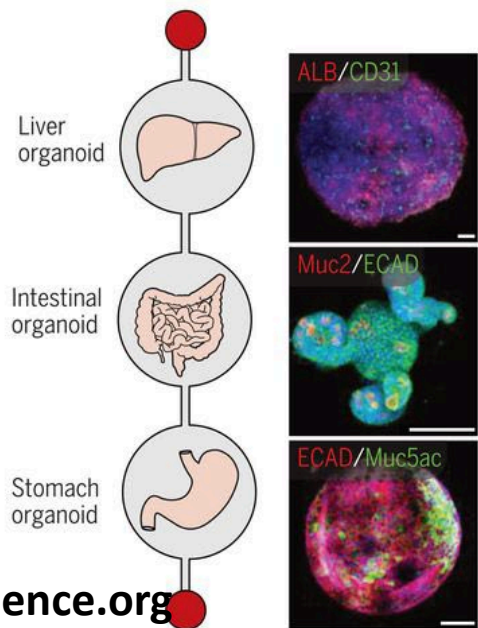
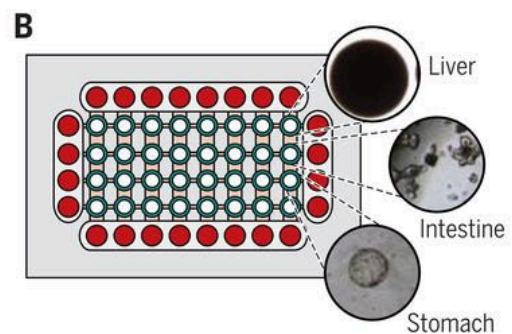
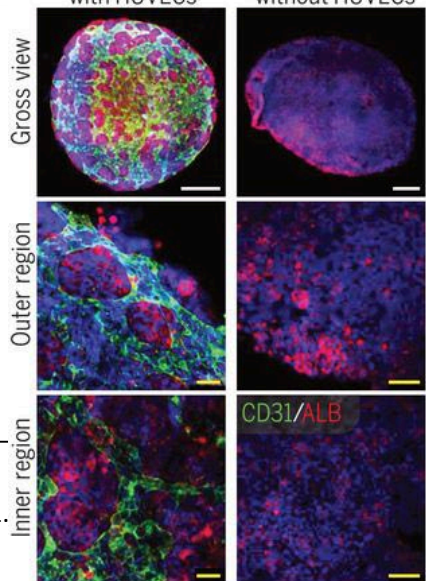
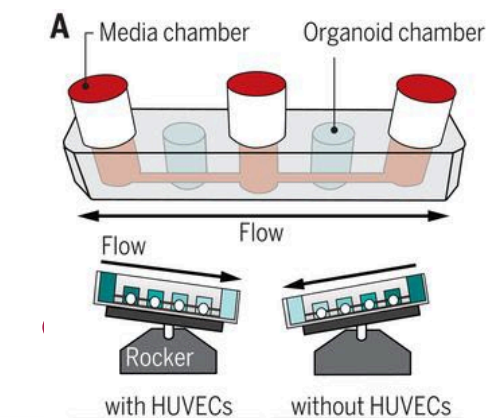
1. „KI-System“ ein maschinengestütztes System, das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können;



## Elemente der Definition in Art. 3 Nr. 1 KI-VO

- 1. Ein maschinengestütztes System ...**
- 2. ... das auf einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist ...**
- 3. ... und das nach seiner Betriebsaufnahme anpassungsfähig sein kann ...**
- 4. ... und das aus den erhaltenen Eingaben ... ableitet, WIE Ausgaben ... erstellt werden ...**
- 5. ... für explizite oder implizite Ziele ...**
- 6. ... Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen ...**
- 7. ... die physische oder virtuelle Umgebungen beeinflussen können.**

# 1. Ein maschinengestütztes System ...



## 1. Ein maschinengestütztes System ...



**Merkmal hilft nicht bei der Abgrenzung von Systemen, von denen KI-Systeme abgegrenzt werden müssten.**

**Im Lichte des Zusammenwachsens von Computer- und Lebenswissenschaften rechtspolitisch zweifelhaft.**

## 2. ... das auf einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist ...

während die von einem KI-System erzeugten Ausgaben verschiedene Funktionen von KI-Systemen widerspiegeln, darunter Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen. KI-Systeme sind mit verschiedenen Graden der Autonomie ausgestattet, was bedeutet, dass sie bis zu einem gewissen Grad unabhängig von menschlichem Zutun agieren und in der Lage sind, ohne menschliches Eingreifen zu arbeiten. Die Anpassungsfähigkeit, die ein KI-System nach Inbetriebnahme aufweisen könnte, bezieht sich auf seine Lernfähigkeit, durch die es sich während seiner Verwendung verändern kann. KI-Systeme können eigenständig oder als Bestandteil eines Produkts verwendet



**AI system autonomy** (contained in both the original and the revised definition of an AI system) means the degree to which a system can learn or act without human involvement following the delegation of autonomy and process automation by humans. Human supervision can occur at any stage of the AI system lifecycle, such as during AI system design, data collection and processing, development, verification, validation, deployment, or operation and monitoring. Some AI systems can generate outputs without these outputs being explicitly described in the AI system's objective and without specific instructions from a human.

## 2. ... das auf einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist ...



**„Bis zu einem gewissen Grad unabhängig von menschlichem Zutun“ operiert wohl fast jedes relevante System!**

**Merkmal hilft nicht bei der Abgrenzung.**

### 3. ... und das nach seiner Betriebsaufnahme anpassungsfähig sein kann ...



**Gilt die KI-VO nur für Systeme, die sich im Betrieb noch weiter anpassen?**

**Und falls ja, was würde das heißen? Nur selbstständige Anpassung des KI-Systems oder auch Nutzung fixierter Modelle mit Updates?**

**KI-VO hätte potenziell (zurzeit) nur einen extrem schmalen Anwendungsbereich!**

### 3. ... und das nach seiner Betriebsaufnahme anpassungsfähig sein kann ...

Autonomie ausgestaltet, was bedeutet, dass sie bis zu einem gewissen Grad unabhängig von menschlichem Zutun agieren und in der Lage sind, ohne menschliches Eingreifen zu arbeiten. Die Anpassungsfähigkeit, die ein KI-System nach Inbetriebnahme aufweisen könnte, bezieht sich auf seine Lernfähigkeit, durch die es sich während seiner Verwendung verändern kann. KI-Systeme können eigenständig oder als Bestandteil eines Produkts verwendet werden, unabhängig davon, ob das System physisch in das Produkt integriert (eingebettet) ist oder der Funktion des



**Adaptiveness** (contained in the revised definition of an AI system) is usually related to AI systems based on machine learning that can continue to evolve after initial development. The system modifies its behaviour through direct interaction with input and data before or after deployment. Examples include a speech recognition system that adapts to an individual's voice or a personalised music recommender system. AI systems can be trained once, periodically, or continually and operate by inferring patterns and relationships in data. Through such training, some AI systems may develop the ability to perform new forms of inference not initially envisioned by their programmers.

## 4. ... und das aus den erhaltenen Eingaben ... ableitet, WIE Ausgaben ... erstellt werden ...

naturalen Personen definierten Regeln für das automatische Ausführen von Operationen beruhen. Ein wesentliches Merkmal von KI-Systemen ist ihre Fähigkeit, abzuleiten. Diese Fähigkeit bezieht sich auf den Prozess der Erzeugung von Ausgaben wie Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen, die physische und digitale Umgebungen beeinflussen können, sowie auf die Fähigkeit von KI-Systemen, Modelle oder Algorithmen oder beides aus Eingaben oder Daten abzuleiten. Zu den Techniken, die während der Gestaltung eines KI-Systems das Ableiten ermöglichen, gehören Ansätze für maschinelles Lernen, wobei aus Daten gelernt wird, wie bestimmte Ziele erreicht werden können, sowie logik- und wissensgestützte Konzepte, wobei aus kodierten Informationen oder symbolischen Darstellungen der zu lösenden Aufgabe abgeleitet wird. Die Fähigkeit eines KI-Systems, abzuleiten, geht über die einfache Datenverarbeitung hinaus, indem Lern-, Schlussfolgerungs- und Modellierungsprozesse ermöglicht werden.



### “INFERRING HOW TO” GENERATE OUTPUTS

The concept of “inference” generally refers to the step in which a system generates an output from its inputs, typically after deployment.<sup>5</sup> When performed during the build phase, inference, in this sense, is often used to evaluate a version of a model, particularly in the machine learning context. In the context of this explanatory memorandum, “infer how to generate outputs” should be understood as also referring to the build phase of the AI system, in which a model is derived from inputs/data.



## 4. ... und das aus den erhaltenen Eingaben ... ableitet, WIE Ausgaben ... erstellt werden ...



**Merkmal und seine Formulierung scheinen im Lichte von ErwGr 12 und OECD völlig missglückt.**

**Es reicht offenbar, wenn**

- **in der Entwicklungsphase aus Trainings-Daten ein Modell etc. abgeleitet wird.**

**ODER**

- **in der Betriebsphase aus den Eingabe-Daten die Ausgaben abgeleitet werden (aber das macht doch jedes System!)**

## 5. ... für explizite oder implizite Ziele ...



**Was ist mit den Zielen gemeint? Die von den Anbietern oder Betreibern gesetzten Ziele?**

**Oder vielleicht zielorientierte Optimierung?  
(zB „Rückwärtsrechnen“ im Vergleich zum „Vorwärtsrechnen“)**

**ErwGr 12 und OECD lassen nicht darauf schließen, dass dem Merkmal entscheidende Abgrenzungsfunktion zukommt.**

## 6. ... Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen ...



**Sinnvolles Merkmal, weil es komplexe Aufgabenstellungen, bei denen es *ex ante* betrachtet nicht das eine, formal determinierte und eindeutig richtige oder falsche Ergebnis gibt, von einfachen Berechnungen abgrenzt.**

## 7. ... die physische oder virtuelle Umgebungen beeinflussen können.



**Merkmal trägt wieder nichts zur Abgrenzung von anderen Systemen bei.**

*Artikel 96*

**Leitlinien der Kommission zur Durchführung dieser Verordnung**

(1) Die Kommission erarbeitet Leitlinien für die praktische Umsetzung dieser Verordnung, die sich insbesondere auf Folgendes beziehen:

[...]

f) die Anwendung der Definition eines KI-Systems gemäß Artikel 3 Nummer 1.

**Fazit: Kommission hat fast völlige Freiheit bei der Konkretisierung!**



## Drei-Faktor-Ansatz

**Faktor I:**

**Rolle von Daten oder  
(domänenspezifischem)  
Erfahrungswissen bei der  
Entwicklung**

**Faktor II:**

**Rolle zielorientierter  
Optimierung bei der  
Anwendung (schließt ggf  
Adaptivität ein)**

**Faktor III:**

**Ausmaß formaler  
Unbestimmtheit bei den  
Ausgaben**

Vgl Wendehorst, Nessler et al in MMR 2024

| Anwendung                                       | Faktor I<br>Entstehung | Faktor II<br>Funktion | Faktor III<br>Ergebnisse/<br>Einsatzfelder | Empfehlung    |
|---|------------------------|-----------------------|--|---------------|
| OCR<br>(maschinengeschriebener Text)            | ++                     | 0                     | 0  | nicht erfasst |
| OCR<br>(Handschrift)                            | ++                     | 0/++                  | +  | erfasst       |
| Chatbot auf Basis eines festen Open Source LLMs | ++                     | 0                     | ++   | erfasst       |
| Chatbot mit API-Zugang zu einem LLM             | ++                     | ++                    | ++   | erfasst       |
| Suchmaschine<br>Google, Duck-Duck-Go, Bing      | ++                     | 0                     | ++   | erfasst       |
| Empfehlungssystem in Amazon                     | ++                     | ++                    | ++   | erfasst       |
| Timeline und Filterung bei X (früher Twitter)   | ++                     | ++                    | ++   | erfasst       |

|   |        |        |        |                 |
|---|--------|--------|--------|-----------------|
| Empfehlungssystem Spotify                                   | ++     | 0      | ++     | erfasst         |
| Bonitäts-Checks (zB KSV und SCHUFA)                         | +/++   | 0      | ++     | erfasst         |
| Medizinische Bild diagnostik (zB Hautkrebs, CT-/MRT-Bilder) | ++     | 0      | ++     | erfasst         |
| Stundenplan-Generator                                       | 0      | +      | +/0    | nicht erfasst   |
| Routenplaner, Navigationssystem                             | 0      | +      | 0      | nicht erfasst   |
| Predictive Maintenance                                      | ++     | 0      | ++     | erfasst         |
| Adaptiver Heizungsregler                                    | 0      | ++     | 0      | nicht erfasst   |
| Adaptive Regler in der Robotik (zB „Spot“)                  | 0      | ++     | +      | erfasst         |
| Excel Sheet   | ++/+/0 | ++/+/0 | ++/+/0 | (nicht) erfasst |
| Rechtschreib- und Grammatikprüfung                          | ++     | 0      | 0      | nicht erfasst   |

Wenn noch Fragen  
offen sind ... ;-)



Ankündigung

**Martini / Wendehorst**  
**KI-VO**

Künstliche Intelligenz-Verordnung

**Kommentar**

Buch. Hardcover

2024

Rund 1200 S.

C.H.BECK. ISBN 978-3-406-81136-4

Format (B x L): 16,0 x 24,0 cm

## Zusammenfassung

- Die Definition eines KI-Systems nach der KI-VO beinhaltet 7 Merkmale, von denen 6 entweder gar nicht auf eine Abgrenzung von anderen Systemen abzielen oder aber aufgrund ihrer Widersprüchlichkeit nicht geeignet sind, eine solche Abgrenzung vorzunehmen.
- Die Kommission hat es mit dieser Definition in der Hand, in ihren Leitlinien den Anwendungsbereich der KI-VO fast beliebig festzulegen – er könnte fast alle Systeme oder aber auch nur die (derzeit extrem wenigen) Systeme erfassen, die sich noch im Betrieb selbständig anpassen.
- Liest man die Definition im Lichte von ErwGr 12 und dem Explanatory Memorandum zur überarbeiteten OECD-Definition, erscheint es als schlüssig, von drei Faktoren auszugehen, welche im Wege eines beweglichen Systems zusammenspielen: (I) Die Rolle von Daten oder Erfahrungswissen bei der Entwicklung; (II) die Rolle von zielorientierter Optimierung bei der Anwendung; und (III) das Ausmaß formaler Unbestimmtheit bei den Ausgaben.