

## „Europa sollte auf dem Gebiet der KI-Ethik eine Vorreiterrolle einnehmen“

**Utl.: ExpertInnengruppe mit Mark Coeckelbergh entwickelt Ethik-Leitlinie für künstliche Intelligenz**

Quelle: [medienportal.univie.ac.at](https://medienportal.univie.ac.at), 4.4.2019

Am Dienstag, 9. April präsentiert die Europäische Kommission Ethik-Leitlinien für eine vertrauenswürdige künstliche Intelligenz (KI). Entwickelt wurden diese Leitlinien von einer Gruppe von ExpertInnen, darunter auch Technikphilosoph Mark Coeckelbergh von der Universität Wien, der ebenso Mitglied im österreichischen Rat für Robotik und Künstliche Intelligenz ist.

Vertrauenswürdige künstliche Intelligenz meint eine KI, die geltenden Gesetzen und Regelungen, sowie ethischen Grundsätzen und Werten folgt und nicht zu unbeabsichtigtem Schaden führt.

„Das Thema der künstlichen Intelligenz-Ethik genießt gerade große Aufmerksamkeit. Die Leitlinien sind allerdings nur ein Dokument, es ist auch an der Zeit zu handeln. Wir müssen uns jetzt um die Regulierung von KI kümmern, da es ansonsten zu spät ist“, stellt Mark Coeckelbergh von der Universität Wien klar, der gemeinsam mit internationalen ExpertInnen an der Entwicklung der Leitlinien gearbeitet hat. Das Rahmenwerk umfasst abstrakte ethische Grundsätze, konkretere Leitlinien und auch Anwendungsbeispiele.

Allgemein empfehlen die ExpertInnen, dass KI-Systeme die ethischen Grundsätze der Achtung der menschlichen Autonomie, der Schadensvermeidung, der Gerechtigkeit und der Erklärbarkeit befolgen sollten. Besondere Aufmerksamkeit gilt dabei schutzbedürftigen Gruppen, wie Kindern, Menschen mit Behinderungen und anderen Gruppen, die historisch benachteiligt wurden oder asymmetrischen Machtverhältnissen ausgesetzt sind.

Konkret führt die Leitlinie sieben Anforderungen an eine vertrauenswürdige KI an: Gestaltung und Kontrolle unter menschlicher Aufsicht, Robustheit und Sicherheit der Technologie, Datenschutz und Datenqualitätsmanagement, Transparenz, Diversität, Nichtdiskriminierung und Fairness, Berücksichtigung gesellschaftlicher und ökologischer Konsequenzen, und Rechenschaftspflicht. Zur Umsetzung dieser Anforderungen schlägt das Rahmenwerk verschiedene Methoden vor, wie die Rückverfolgbarkeit und Überprüfbarkeit von KI-Systemen, Forschung, die Diskussion von Fragen in der Öffentlichkeit und die Einbeziehung unterschiedlicher Stakeholder in Entwurf und Entwicklung von KI-Systemen.

Bei der Bewertung von vertrauenswürdiger KI gehe es laut Leitlinie nicht um das „Abhaken von Kästchen“. Vielmehr sollen, „fortlaufend Anforderungen identifiziert und umgesetzt, Lösungen evaluiert und bessere Resultate während des gesamten Lebenszyklus des KI-Systems erzielt werden, dabei sind auch unterschiedliche Stakeholder in den Prozess miteinzubeziehen“. Für den Technikphilosophen ist insbesondere der Begriff „gemeinsam“ in diesem Zusammenhang wichtig: „Was zählt, ist, dass die KI den Menschen und der Gesellschaft nutzt. Die BürgerInnen sollen von künstlicher Intelligenz profitieren, und zwar möglichst alle BürgerInnen und nicht nur diejenigen, die ohnehin schon begünstigt sind.“

Aus ExpertInnensicht sollen diese Leitlinien keineswegs als Ersatz für politische Entscheidungen dienen, sondern stellen viel-

mehr einen Ausgangspunkt für Diskussionen über vertrauenswürdige KI in Europa dar und zielen – über Europa hinaus – auf die Förderung von Forschung, Reflexion und Diskussion über KI-Ethik auf globaler Ebene ab. „Wenn es Europa gelingt, ethische Aspekte in die Entwicklung von KI-Innovationen miteinzubeziehen, wird Europa eine Vorbildwirkung für die ganze Welt haben, in der Akteure wie die Vereinigten Staaten und China bereits verstärkt auf das Vorgehen der EU blicken“, so Coeckelbergh.

### Wissenschaftlicher Kontakt

Univ.-Prof. Dr. Mark Coeckelbergh  
Institut für Philosophie  
Universität Wien  
1010 - Wien, Universitätsstraße 7  
+43-1-4277-464 66  
+43-664-602 77-464 66  
[mark.coeckelbergh@univie.ac.at](mailto:mark.coeckelbergh@univie.ac.at)

### Rückfragehinweis

Paulina Parvanov, BA MA  
Pressebüro  
Universität Wien  
1010 - Wien, Universitätsring 1  
+43-1-4277-175 40  
+43-664-60277-175 40  
[paulina.parvanov@univie.ac.at](mailto:paulina.parvanov@univie.ac.at)