

13/2020

02. April 2020

## Neue Studie von VDE und Bertelsmann Stiftung:

### **KI-Ethik messbar machen**

- **In einer neuen Studie zeigen VDE und Bertelsmann Stiftung, wie Ethik in der KI konkret umsetzbar und messbar wird**
- **„Hands-on“: Studie gibt Unternehmen Werkzeug in die Hand, KI anzuwenden und liefert notwendige Orientierung**
- **Modell international anwendbar: VDE und Bertelsmann Stiftung wollen KI „made in Europe“ zum Durchbruch verhelfen**

(Frankfurt a.M./Gütersloh, 02.04.2020) In einer neuen Studie zeigen die Technologieorganisation VDE und die Bertelsmann Stiftung, wie sich Ethikprinzipien für Künstliche Intelligenz (KI) in die Praxis bringen lassen. Zwar gibt es eine Vielzahl an Initiativen, die ethische Richtlinien für die Gestaltung von KI veröffentlicht haben, allerdings bis dato kaum Lösungen für deren praktische Umsetzung. Genau hier setzt der VDE als Initiator und Leiter der „AI Ethics Impact Group“ gemeinsam mit der Bertelsmann Stiftung an. Die Studie „From principles to practice – an interdisciplinary framework to operationalise AI ethics“ zeigt, wie mit der geschickten Kombination dreier Werkzeuge – einem WKIO-Modell, einem KI-Ethik-Label und einer Risikoklassifizierung – allgemeine ethische Prinzipien messbar und konkret umsetzbar werden. Das größte Hindernis für die Entwicklung von ethischer KI sind laut VDE und Bertelsmann Stiftung die Unschärfen und die unterschiedlichen Verständnisse von Prinzipien wie „Transparenz“ und „Gerechtigkeit“. Dies führt dazu, dass KI-entwickelnden Unternehmen die nötige Orientierung fehlt und eine wirksame Durchsetzung von KI-Ethik bisher schwierig war.

### **Messbarkeit und Sichtbarkeit durch WKIO-Modell und Ethik-Label**

Das sogenannte WKIO-Modell (im Englischen VCIO: Value, Criteria, Indicators, Observables) schlüsselt Werte in Kriterien, Indikatoren und letztlich messbare Observablen auf. Demonstriert wird dies in der Studie für die Werte Transparenz, Gerechtigkeit und Verantwortlichkeit. Das WKIO-Modell kann von Politikentwicklern, Regulierern und Aufsichtsbehörden genutzt werden, um Anforderungen an KI-Systeme zu konkretisieren und

durchzusetzen. Das ebenfalls im Rahmen der Studie entwickelte Ethik-Label für KI-Systeme ermöglicht Unternehmen, die ethischen Eigenschaften ihrer Produkte klar und einheitlich zu kommunizieren. Das Label lehnt sich an das erfolgreiche Energieeffizienzlabel für Elektrogeräte an und schafft sowohl für Konsumenten als auch Unternehmen eine bessere Vergleichbarkeit der auf dem Markt zur Verfügung stehenden Produkte. „Mit dem Label legen wir die Grundlage für das Setzen regulatorischer Grenzen und schaffen gleichzeitig die Voraussetzungen für einen funktionierenden Markt,“ erklärt Dr. Sebastian Hallensleben vom VDE. Es handele sich nicht um ein ungenaues ja/nein-Gütesiegel, sondern um eine abgestufte Kennzeichnung relevanter Eigenschaften eines Systems, macht Hallensleben deutlich. Welche Eigenschaften sich dabei beispielsweise hinter „Transparenz: Klasse C“ oder „Gerechtigkeit: Klasse B“ verbergen, wird anhand des WKIO-Modells spezifiziert. „Für KI-entwickelnde Unternehmen bietet der Ansatz eine Möglichkeit, ethische Selbstverpflichtungen auszudifferenzieren und ihren Mitarbeitenden konkrete Orientierung für die Umsetzung der darin enthaltenen Prinzipien zu bieten“, sagt Carla Hustedt von der Bertelsmann Stiftung. Organisationen, die KI nutzen wollen, könnten das Modell auch dafür verwenden, Kriterien für die Beschaffung passender KI-Systeme zu definieren.

### **Risiko-Matrix hilft bei Klassifizierung von KI-Anwendungsfällen**

Welche Eigenschaften eines KI-Systems als „ethisch ausreichend“ gelten, ist vom konkreten Anwendungsfall abhängig. Der Einsatz eines KI-Systems zur Verteilung von Gütern dürfte beispielsweise in der Möbellogistik ethisch weniger heikel sein als bei der Belieferung von Krankenhäusern. Die Studie stellt daher mit der „Risiko-Matrix“ einen Ansatz zur Klassifizierung des Anwendungskontexts vor. Während die niedrigste Klasse 0 keine weiteren ethischen Überlegungen erfordert, dürfen für die Klassen 1 bis 3 nur KI-Systeme eingesetzt werden, die das obige KI-Ethik-Label tragen und innerhalb dieses Labels bestimmte Mindestanforderungen erfüllen. Die Einstufung in Klasse 0, 1, 2 oder 3 berücksichtigt einerseits die Intensität des potenziellen Schadens, beispielsweise das Ausmaß der Verletzung von individuellen Grundrechten, die Anzahl der betroffenen Personen oder mögliche negative Auswirkungen auf die Gesellschaft als Ganzes, und andererseits die Abhängigkeit der betroffenen Personen bzw. deren Möglichkeit, sich einer KI-Entscheidung zu entziehen, ein anderes System wählen oder eine bestehende Entscheidung hinterfragen zu können.

Die Studie „From principles to practice – an interdisciplinary framework to operationalise AI ethics“ entstand unter der Federführung von VDE und Bertelsmann Stiftung in der „AI Ethics Impact Group“ und ist kostenlos unter [www.vde.com/presse](http://www.vde.com/presse) bzw. [www.ai-ethics-impact.org](http://www.ai-ethics-impact.org) erhältlich.

### **Über die AI Ethics Impact Group:**

Der VDE und die Bertelsmann Stiftung haben im Herbst 2019 die interdisziplinäre AI Ethics Impact Group ins Leben gerufen und gemeinsam die Studie „AI Ethics: From Principles to Practice - An interdisciplinary framework to operationalise AI ethics“ erarbeitet. In der AI Ethics Impact Group kommen ExpertInnen aus den Bereichen Informatik, Philosophie, Ingenieurs- und Sozialwissenschaften zusammen, insbesondere aus dem Algorithmic Accountability Lab der TU Kaiserslautern, aus dem Höchstleistungsrechenzentrum der Uni Stuttgart, dem Institut für Technikfolgenabschätzung und Systemanalyse (ITAS) in Karlsruhe, dem Institut für Philosophie der TU Darmstadt, dem Internationalen Zentrum für Ethik in den Wissenschaften (IZEW) der Uni Tübingen und dem Thinktank iRights.Lab. Nähere Informationen unter [www.ai-ethics-impact.org](http://www.ai-ethics-impact.org).

### **Über den VDE:**

Der VDE, eine der größten Technologie-Organisationen Europas, steht seit mehr als 125 Jahren für Innovation und technologischen Fortschritt. Als einzige Organisation weltweit vereint der VDE dabei Wissenschaft, Standardisierung, Prüfung, Zertifizierung und Anwendungsberatung unter einem Dach. Das VDE-Zeichen gilt seit 100 Jahren als Synonym für höchste Sicherheitsstandards und Verbraucherschutz. Wir setzen uns ein für die Forschungs- und Nachwuchsförderung und für das lebenslange Lernen mit Weiterbildungsangeboten „on the job“. 2.000 Mitarbeiter an über 60 Standorten weltweit, mehr als 100.000 ehrenamtliche Experten und rund 1.500 Unternehmen gestalten im Netzwerk VDE eine lebenswerte Zukunft: vernetzt, digital, elektrisch. Wir gestalten die e-diale Zukunft. Hauptsitz des VDE (Verband der Elektrotechnik Elektronik und Informationstechnik e.V.) ist Frankfurt am Main. [www.vde.com](http://www.vde.com)

### **Über die Bertelsmann Stiftung:**

Die Bertelsmann Stiftung ist eine der größten operativen Stiftungen in Europa. Sie ist eine gemeinnützige Organisation, die sich in der Tradition ihres Gründers Reinhard Mohn für eine zukunftsfähige Gesellschaft engagiert. Das Projekt „[Ethik der Algorithmen](#)“, im Rahmen dessen die Studie entstanden ist, setzt sich für drei Ziele ein: 1. Effektive Kontrolle algorithmischer Systeme sicherstellen, 2. Vielfalt der Systeme, Betreiber und Akteure fördern und 3. Chancen algorithmischer Systeme für das Gemeinwohl heben. In dem Projekt sensibilisieren wir die Bevölkerung, Politik und Stakeholder aus Wirtschaft und Zivilgesellschaft für Chancen, Risiken und vor allem die Relevanz algorithmischer Prozesse, bieten Foren für internationale, interdisziplinäre und sektorübergreifenden Diskurse und pilotieren konkrete Lösungsansätze für eine teilhabeförderliche Gestaltung der digitalen Sphäre.

**Pressekontakt:** Melanie Unseld, Tel. 069 6308461, [melanie.unseld@vde.com](mailto:melanie.unseld@vde.com)