

PRESSEINFORMATION

PRESSEINFORMATION2. Juli 2019 || Seite 1 | 4

Vertrauen schaffen in Künstliche Intelligenz – Interdisziplinäres Team aus Informatik, Philosophie und Recht definiert Handlungsfelder zur Zertifizierung von KI

Das Zeitalter der Digitalisierung durchdringt fast alle Lebensbereiche – von der Arbeitswelt über den Straßenverkehr und den Gesundheitssektor bis hin zur Kommunikation. Wie kann man sicherstellen, dass besonders Entwicklungen auf Basis von Künstlicher Intelligenz (KI) technisch, ethisch und rechtlich vertretbar sind? Wissenschaftlerinnen und Wissenschaftler der Universitäten Bonn und Köln erarbeiten dazu unter der Federführung des Fraunhofer-Instituts für Intelligente Analyse- und Informationssysteme IAIS und unter Mitwirkung des Bundesamtes für Sicherheit in der Informationstechnik (BSI) einen Prüfkatalog zur Zertifizierung von KI-Anwendungen. Das interdisziplinäre Team stellt die Handlungsfelder aus philosophischer, ethischer, rechtlicher und technologischer Sicht nun in einer Publikation vor.

Künstliche Intelligenz verändert unsere Gesellschaft, unsere Wirtschaft und unseren Alltag grundlegend und eröffnet große Chancen für unser Zusammenleben. Sie hilft zum Beispiel Ärzten, Röntgenbilder besser und oft exakter auszuwerten, beantwortet in Form von Chatbots Fragen zu Versicherungspolice und anderen Produkten und wird in absehbarer Zeit Autos immer selbstständiger fahren lassen. Aktuelle Prognosen gehen davon aus, dass die Zahl der KI-Anwendungen in den nächsten Jahren exponentiell wachsen wird. McKinsey zum Beispiel prognostiziert bis 2030 global bis zu 13 Billionen Dollar zusätzliche Wertsteigerung durch Künstliche Intelligenz.

Gleichzeitig wird deutlich, dass eine sorgfältige Gestaltung von KI-Anwendungen notwendig ist, damit wir die Chancen der KI im Einklang mit unseren gesellschaftlichen Werten und Vorstellungen nutzen können. Deshalb hat sich das interdisziplinäre Expertenteam im Kontext der Kompetenzplattform KI.NRW zusammengetan, um eine durch akkreditierte Prüfer durchzuführende Zertifizierung für KI-Anwendungen zu entwickeln. Das Zertifikat soll einen Qualitätsstandard bescheinigen, der es Technologieanbietern erlaubt, KI-Anwendungen überprüfbar technisch zuverlässig und ethisch akzeptabel zu gestalten.

»Mit der Zertifizierung wollen wir dazu beitragen, Qualitätsstandards für eine ‚KI Made in Europe‘ zu setzen, den verantwortungsvollen Umgang mit der Technologie zu sichern und einen fairen Wettbewerb verschiedener Anbieter zu befördern«, sagt Prof. Dr.

Redaktion

Katrin Berkler M.A. | Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS | Telefon +49 2241 14-2252 | Schloss Birlinghoven | 53757 Sankt Augustin | www.iais.fraunhofer.de | pr@iais.fraunhofer.de |

Stefan Wrobel, Institutsleiter des Fraunhofer IAIS und Professor für Informatik an der Universität Bonn.

PRESSEINFORMATION2. Juli 2019 || Seite 2 | 4

Der Mensch im Mittelpunkt

Künstliche Intelligenz hat das Potenzial, die Fähigkeiten des Menschen zu erweitern und hilft uns, neue Erkenntnisse zu gewinnen. Wenn wir unseren Entscheidungen jedoch Ergebnisse von voll- oder teilautomatisierten selbstlernenden Systemen zugrunde legen, stellt uns dies gleichzeitig vor neue Herausforderungen. Neben der technischen Eignung müssen vor allem grundsätzliche philosophisch-ethische Aspekte, aber auch rechtliche Fragestellungen geklärt werden. Um sicherzustellen, dass der Mensch stets im Mittelpunkt dieser Entwicklung steht, ist daher ein enger Austausch zwischen Informatik, Philosophie und Rechtswissenschaften notwendig.

In dem nun veröffentlichten »Whitepaper« erläutern die Expertinnen und Experten ihre interdisziplinäre Herangehensweise an die KI-Zertifizierung. So gehen die Autorinnen und Autoren etwa auf die ethischen Grundsätze ein: »Alle Beteiligten sollen sich beim Einsatz von KI nach ihren moralischen Überzeugungen gut verhalten können, und niemand darf in seinen Rechten, seiner Autonomie oder Freiheit beschnitten werden«, betont Prof. Dr. Markus Gabriel, Professor für Philosophie an der Universität Bonn. Auch juristische Fragestellungen werden adressiert: »Es muss zum Beispiel geklärt werden, welche Anforderungen an KI-Anwendungen sich aus den Grundwerten und Prinzipien eines freiheitlich geordneten Rechtsstaates ergeben«, sagt Prof. Dr. Dr. Frauke Rostalski, Professorin für Rechtswissenschaft an der Universität zu Köln.

Handlungsfelder für den vertrauensvollen Einsatz von KI

Aus diesem interdisziplinären Ansatz gehen mehrere Handlungsfelder aus ethischer, rechtlicher und technologischer Sicht hervor, die im Whitepaper beleuchtet werden. Kriterien für die Zertifizierung sollen Fairness, Transparenz, Autonomie und Kontrolle, Datenschutz sowie Sicherheit und Verlässlichkeit sein. Damit orientiert sich das Zertifizierungsprojekt von KI.NRW auch an Empfehlungen der EU.

Typische Fragen innerhalb der Zertifizierung sollen sein: Respektiert die KI-Anwendung gesellschaftliche Werte und Gesetze? Ist eine selbstbestimmte, effektive Nutzung möglich? Behandelt die KI alle Betroffenen fair? Sind ihre Funktionsweise und Entscheidungen nachvollziehbar? Funktioniert die KI zuverlässig und ist sie robust? Ist sie sicher gegen Angriffe, Unfälle und Fehler? Schützt die KI die Privatsphäre und sonstige sensiblen Informationen?

Redaktion

Katrin Berkler M.A. | Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS | Telefon +49 2241 14-2252 | Schloss Birlinghoven | 53757 Sankt Augustin | www.iais.fraunhofer.de | pr@iais.fraunhofer.de |

Leitplanken bereits beim Design

PRESSEINFORMATION2. Juli 2019 || Seite 3 | 4

Bereits während des Designs der Technologien müsse geklärt werden, ob die Anwendung ethisch und rechtlich zulässig ist und falls ja, welche Leitplanken für ihre Ausgestaltung formuliert werden sollen, so die Autorinnen und Autoren. Ein notwendiges Kriterium sei, allen Beteiligten dieselben Möglichkeiten einer moralischen Entscheidung zu geben, die sie auch im Falle eines Verzichts auf den KI-Einsatz hätten, und ihre Rechte sowie ihre Freiheit zu achten.

Ein wichtiges Kriterium ist auch die Transparenz. Informationen zum richtigen Umgang mit der KI-Anwendung müssten verfügbar sein, betonen die Experten. Die Interpretierbarkeit, Nachverfolgbarkeit und Reproduzierbarkeit von Ergebnissen der Künstlichen Intelligenz müssten für die Nutzer möglich sein. Widerstreitende Interessen – wie Transparenz und Geheimhaltung von Geschäftsgeheimnissen – müssten gegeneinander abgewogen werden.

Es ist geplant, bis Anfang 2020 eine erste Version des Prüfkatalogs zu veröffentlichen und die ersten KI-Anwendungen zu zertifizieren. Dazu kooperieren die Projektverantwortlichen mit dem Bundesamt für Sicherheit in der Informationstechnik (BSI), dessen langjährige Erfahrung in der Entwicklung von sicheren IT-Standards in die Zertifizierung einfließt. Da sich die Künstliche Intelligenz ständig weiterentwickelt, wird auch der Prüfkatalog ein »lebendes Dokument« sein, das ständiger Aktualisierung bedarf.

Beteiligte Institutionen

[Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS](#) (Institutsleiter Prof. Dr. Stefan Wrobel, Projektleiter Dr. Maximilian Poretschkin); [Universität Bonn, Center for Science and Thought](#) (Direktor Prof. Dr. Markus Gabriel, Projektleiter Jan Voosholz); [Universität zu Köln, Lehrstuhl für Strafrecht, Strafprozessrecht, Rechtsphilosophie und Rechtsvergleichung](#) (Inhaberin Prof. Dr. Dr. Frauke Rostalski, Projektleitung).

Download Whitepaper

Das Whitepaper steht als PDF zum Download bereit unter:

www.iais.fraunhofer.de/ki-zertifizierung

Weitere Informationen

Kompetenzplattform KI.NRW: www.ki.nrw

Redaktion

Katrin Berkler M.A. | Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS | Telefon +49 2241 14-2252 | Schloss Birlinghoven | 53757 Sankt Augustin | www.iais.fraunhofer.de | pr@iais.fraunhofer.de |

FRAUNHOFER-INSTITUT FÜR INTELLIGENTE ANALYSE- UND INFORMATIONSSYSTEME IAIS

Ansprechpartner*innen

Pressekontakt

Katrin Berkler
Telefon 02241 14-2252
pr@iais.fraunhofer.de

Projektleitung Fraunhofer IAIS

Dr. Maximilian Poretschkin
Telefon 02241 14-1984
maximilian.poretschkin@iais.fraunhofer.de

Projektleitung Universität Bonn

Jan Voosholz
Telefon 0228 73-62042
voosholz@uni-bonn.de

Projektleitung Universität zu Köln

Prof. Dr. Dr. Frauke Rostalski
Telefon 0221 470-4284
ls-rostalski@uni-koeln.de

PRESSEINFORMATION

2. Juli 2019 || Seite 4 | 4

Redaktion

Katrin Berkler M.A. | Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS | Telefon +49 2241 14-2252 |
Schloss Birlinghoven | 53757 Sankt Augustin | www.iais.fraunhofer.de | pr@iais.fraunhofer.de |