

CERTAIN – Ein Zentrum für vertrauenswürdige Künstliche Intelligenz

Das **Center for European Research in Trusted AI (CERTAIN)** ist eine Initiative des Deutschen Forschungszentrums für Künstliche Intelligenz (DFKI). Ziel ist es, vertrauenswürdige Künstliche Intelligenz (KI) in Europa konkret zu gestalten – durch wissenschaftlich fundierte, praktisch anwendbare und gesellschaftlich akzeptierte Zertifizierungs- und Prüfverfahren für KI-Systeme. CERTAIN bringt dazu Akteure aus Wissenschaft, Wirtschaft, Politik und Zivilgesellschaft in einem offenen Konsortium zusammen, um gemeinsam an Standards, Methoden und Tools für „Trusted AI“ zu arbeiten.

Im Mittelpunkt steht die Entwicklung von Garantien über das Verhalten von KI – vor allem als integraler Bestandteil des Systemdesigns anstelle eines rein technischen Prüfmechanismus. CERTAIN verfolgt hierbei einen umfassenden Ansatz: Vertrauenswürdigkeit soll bereits im Design angelegt sein, durch robuste, nachvollziehbare und faire Systemarchitekturen. Ergänzend werden technische Verfahren zur Verifikation und Validierung entwickelt, die belastbare Aussagen über Leistungsfähigkeit, Sicherheit und Grenzen ermöglichen. Transparenz wird als zentrale Voraussetzung für Rechenschaft und Nachvollziehbarkeit verstanden – nicht nur durch Offenlegung, sondern durch echte Erklärbarkeit. Darüber hinaus setzt CERTAIN auf menschenzentrierte Interaktion: Nutzer sollen befähigt werden, KI-Systeme zu verstehen, einzuordnen und bei Bedarf korrigierend einzugreifen.

Diese mehrdimensionale Herangehensweise macht deutlich: CERTAIN schafft keine Etiketten, sondern echte Orientierung. Ziel ist es, Vertrauen in KI nicht vorauszusetzen, sondern durch überprüfbare Kriterien systematisch herzustellen – technisch fundiert, institutionell abgesichert und kommunizierbar.

Im Zuge der regulatorischen Anforderungen der europäischen KI-Verordnung (EU AI Act) arbeitet CERTAIN an der Entwicklung eines modularen Bildungsangebots zur vertrauenswürdigen Anwendung von KI. Ziel ist es, zentrale Anforderungen wie Risikomanagement, Transparenz, Fairness und technische Robustheit praxisnah zu vermitteln – insbesondere für Akteure, die KI-Systeme entwickeln, betreiben oder bewerten. Die Schulungsmodule sollen zukünftig eine Grundlage für qualifizierte Weiterbildungen und zertifizierbare Kompetenznachweise bieten – und so zur Umsetzung des EU AI Act in der Breite beitragen.

CERTAIN agiert transdisziplinär und vernetzt. Die Initiative vereint Expertise aus Informatik, Rechtswissenschaft, Ethik, Governance und Normungsarbeit. Gleichzeitig verfolgt CERTAIN einen modernen Ansatz in der Wissenschaftskommunikation: komplexe Inhalte werden klar, zugänglich und visuell ansprechend vermittelt – um Vertrauen nicht nur technisch, sondern auch gesellschaftlich zu verankern.

Mit Sitz im Saarland leistet CERTAIN einen Beitrag zur strategischen Positionierung Deutschlands und Europas als Standort für sichere und verantwortungsvolle KI. Das klare Ziel: Trusted AI – nicht nur versprochen, sondern garantiert.

Kontakt

Für weitere Informationen und Anfragen steht das CERTAIN-Team unter der folgenden Adresse zur Verfügung:

CERTAIN – Centre for European Research in Trusted AI

Stuhlsatzenhausweg 3
Saarland Informatics Campus
Gebäude D 3_2
66123 Saarbrücken
Deutschland

Telefon: +49 681 857 75 54 81

E-Mail: info@certain-trust.eu

Website: www.certain-trust.eu

Pressekontakt

Für Presseanfragen wenden Sie sich bitte an info@certain-trust.eu.

Über das DFKI

Das Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI) ist eine der weltweit führenden Forschungseinrichtungen auf dem Gebiet der Künstlichen Intelligenz. Mit Standorten in Kaiserslautern, Saarbrücken, Bremen, Osnabrück, Berlin, Darmstadt und Lübeck deckt das DFKI ein breites Spektrum an Forschungsthemen ab und arbeitet eng mit Industrie, Wissenschaft und Gesellschaft zusammen.

DFKI – Deutsches Forschungszentrum für Künstliche Intelligenz GmbH

Trippstadter Str. 122

67663 Kaiserslautern

Deutschland

Telefon: +49 631 20575 0

