

Agenda

Virtueller Workshop „Vertrauenswürdige KI“ am 13. Oktober 2021, organisiert vom Fraunhofer ISI

Teil 1 - Umsetzung

Wo stehen wir bei der praktischen Umsetzung einer vertrauenswürdigen KI?

Impulsgeber:

Dr. Thilo Hagendorff (Uni Tübingen)

Dr. Christian Winkler (Datenizing)

Leitfragen für die Diskussion

- Woran liegt es, dass es gut ausformulierte Anforderungen und Kriterienkataloge für die vertrauenswürdige KI gibt, aber noch zu wenige konkrete Beispiele von Umsetzungen in Unternehmen und Organisationen vorliegen?
- Sind die vorhandenen Guidelines nicht praktisch genug?
- Stimmt dieser Eindruck überhaupt oder halten sich Unternehmen und Organisationen nur "bedeckt" und arbeiten bereits aktiv mit vorhandenen oder eigenen Guidelines?
- Wie könnte man zu einer Beispielsammlung kommen, in der Ansätze und Erfahrungen dokumentiert sind und an der sich andere orientieren könnten?

Teil 2 - Strategien

Welche Gründe gibt es für das Umsetzungsdefizit?

Impulsgeber:innen:

Lajla Fetic (Bertelsmann Stiftung)

Dr. Christoph Peylo (Bosch)

Leitfragen für die Diskussion

- Welche Strategien eignen sich, um zu einer größeren Zahl von Umsetzungen vertrauenswürdiger KI in Unternehmen und Organisationen zu kommen?
- Vorschläge: Konkretere Leitfäden; Klärung der Frage, für welche KI-Anwendungen welche Aspekte der vertrauenswürdigen KI überhaupt relevant sind; freiwillige und verpflichtende Zertifizierungen, Regulierung usw.

Teil 3 - Partizipation

Welchen Beitrag kann die partizipative Technikgestaltung leisten?

Impulsgeber:

Dr. Andreas Bischof (Uni Chemnitz)

Leitfragen für die Diskussion

- Vertrauenswürdige KI erfordert einen interdisziplinären Ansatz, bei dem Informatiker:innen, Manager:innen, Ethiker:innen und Sozialwissenschaftler:innen zusammenarbeiten müssen. In welcher Weise kann man hier von den Erfahrungen der partizipativen Technikforschung profitieren und welche Methoden bieten sich an?
- Vorschläge sind: „Embedded“ Social Scientists, ganzheitliches Design nach dem Vorbild von Privacy-by-Design sowie Co-Creation-Ansätze.