

## Online-Veranstaltung „vertrauenswürdige KI mit IEEE“

### Zusammenfassung der Vorträge vom 02.02.2022

Die aws und die Plattform Industrie 4.0 haben eine Initiative gestartet, um österreichische Unternehmen zur Teilnahme an Standardisierungsgremien über vertrauenswürdige KI zu motivieren. Mit regelmäßigen Veranstaltungen über dieses wichtige Thema sollen Unternehmen die Möglichkeit bekommen, mehr über die Gremienarbeit und die Wichtigkeit von Standards zu erfahren. Diesmal gab es Unterstützung durch die [IEEE](#), der internationalen Standardisierungsorganisation, die bereits zahlreiche Standards im Bereich vertrauenswürdiger KI geschaffen hat.

Clara Neppel, Senior Director vom Wiener Standort der IEEE stellte in ihrem Vortrag die IEEE Initiative Ethically Aligned Design vor, wo Technologen gemeinsam mit Stakeholdern der Gesellschaft Lösungsvorschläge für die Aufnahme von ethischen Werten ins Design ausarbeiten. Daraus resultierten bei der IEEE 2016 die ersten Standardentwicklungen. Fünf Jahre später, 2021, wurde der 7000er Standard veröffentlicht. Aktuell arbeitet die IEEE mit der Europäischen Kommission zusammen und untersucht bestehende Standards, die für den [EU AI Act](#) angewendet werden können. Folgende Maßnahmen der IEEE unterstützen die vertrauenswürdige KI: Technische und soziotechnische Standards, ethische Zertifizierung und Governance (Unternehmens-/Organisationskultur).

		
<p><b>STANDARDS</b></p>	<p><b>CERTIFICATION</b></p>	<p><b>GOVERNANCE</b></p>
<p>Developing a growing series of standards that promote innovation, foster interoperability and honor human values.</p> <ul style="list-style-type: none"> <li>- <b>Technical</b></li> <li>- <b>Socio-technical</b></li> </ul>	<p>Developing metrics and processes towards the implementation of a certification methodology.</p> <ul style="list-style-type: none"> <li>- <b>Transparency</b></li> <li>- <b>Accountability</b></li> <li>- <b>Algorithmic bias</b></li> </ul>	<p>Support responsible Artificial Intelligence Systems innovation through proper governance mechanisms.</p> <ul style="list-style-type: none"> <li>- <b>Business</b></li> <li>- <b>Cities</b></li> <li>- <b>Public bodies</b></li> </ul>
<p><a href="https://standards.ieee.org/initiatives/artificial-intelligence-systems">https://standards.ieee.org/initiatives/artificial-intelligence-systems</a> </p>		

Die IEEE hat früh begonnen, die ethischen Prinzipien der EU in ihre Standards zu etablieren. Der 7000er Standard wurde letztes Jahr veröffentlicht. Mehr Informationen unter <https://standards.ieee.org/ieee/7000/6781/>.

Wie ein Programm von KundInnenanfragen ethisch zertifiziert werden kann, haben die [Wiener Stadtwerke](#) mit der IEEE als Kooperationspartner gezeigt. Dietmar Schabus von den Wiener Stadtwerken präsentierte die intern entwickelte KI-Lösung, die automatisch schriftliche Anfragen der Wien Energie-KundInnen kategorisiert und die die ethische Auswertung und Designation durchlaufen hat. Im Zertifizierungsprozess wurde die KI-Lösung einer Risikobewertung durch die IEEE unterzogen, wo das Endergebnis eine Einstufung in Overall-Low-Risk, also eine niedrige Risikoeinstufung, war. 43 ethische Kriterien, die von der IEEE zusammengestellt wurden, mussten durch die KI-Lösung erfüllt werden. Infolgedessen wurde die KI-Lösung durch die IEEE als ethisch vertretbar ausgewiesen.

Beispiele für ethische Kriterien:

Technische Kriterien:

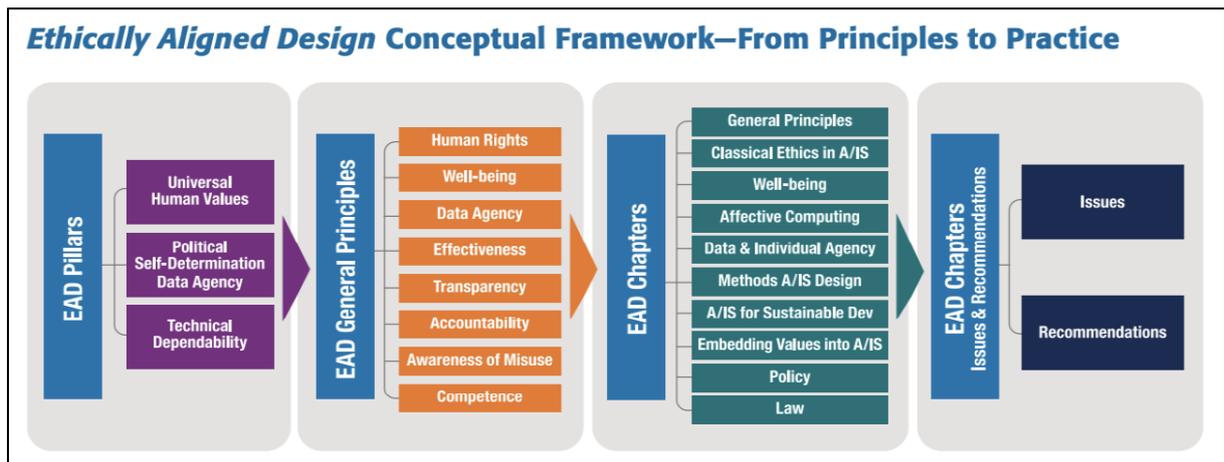
- Error Analysis
- Hyperparameter Tuning
- Mitigate False Positives/Alarms

Organisatorische Kriterien:

- Adopting a Layered Approach
- (Avoid) Inaction, Delay and Indifference
- Human Authority & Autonomy

Alexandra Ebert von [Mostly AI](#), einem Unternehmen zur Erstellung von synthetischen Daten mittels KI, sprach zu dem großen Thema Datenschutz und die Einsatzmöglichkeiten von synthetischen Daten. Durch das zunehmende Wachstum an Datenmengen (Big Data Space) gewinnen Regularien immer mehr an Bedeutung, sowohl auf nationaler als auch auf europäischer/internationaler Ebene, und der Zugriff zu den Daten fürs KI-Training wird damit stark eingeschränkt. Werden die Daten anonymisiert und können sie nicht mehr reidentifiziert werden, fallen sie nicht mehr unter den Anwendungsbereich der DSGVO und können somit verwendet und verarbeitet werden. Das bisherige Problem aber ist, dass nach der Anwendung aktueller Anonymisierungstechnologien die Daten kaum nutzbar sind. Eine Lösung dafür bieten synthetische Daten, welche das Ergebnis einer neuartigen Anonymisierungstechnik sind. Durch einen synthetischen Datengenerator, welcher mit Echtdaten lernt, kann Datenmaterial erzeugt werden, das keine Rückschlüsse auf ein echtes Individuum zulässt. Mostly AI bietet seinen Kunden eine Software zur Generierung von synthetischen Daten (z.B. Bank- oder Versicherungsdaten) basierend auf Echtdaten an. Seine Kunden können die synthetischen Datensätze datenschutzfreundlich nutzen, teilen und so Innovationen fördern.

Vom Hauptsitz der IEEE in New Jersey meldete sich John Havens, Director Emerging Technology & Strategic Development, der aufzeigte, wie ethische Fragestellungen in Entwicklungsprozesse integriert werden können und präsentierte dabei das IEEE Paper [A Call to Action for Businesses Using AI](#). Die kürzlich gegründete AI Ethics Initiative, welche zur IEEE Standards Association gehört, hat das Paper A call to Action for Businesses Using AI publiziert. Es beschreibt, was verantwortungsvolle KI bedeutet, was die Prinzipien des Ethically Aligned Designs (siehe Grafik) sind, wie man eine nachhaltige Kultur schafft uvm.



IEEE, Ethically Aligned Design

Die Präsentationsfolien der Vortragenden finden Sie unter <https://plattformindustrie40.at/services/#vertrauen>

#### Förderinitiative der aws und Plattform Industrie 4.0

Im Rahmen des Förderprogramms „aws Digitalisierung – vertrauenswürdige künstliche Intelligenz“ arbeiten die Plattform Industrie 4.0 Österreich und die aws zusammen: Die Initiative bietet Unternehmen Hilfestellungen bei der Entwicklung von Normen und Standardisierungen für KI-basierte Produkte, Prozesse und Dienstleistungen. Dabei wird besonderer Wert auf vertrauenswürdige Aspekte bei der Entwicklung und dem Einsatz von KI gelegt. Weitere Informationen und die Leitlinien der finanziellen Unterstützung finden Sie unter <https://plattformindustrie40.at/services/#vertrauen>.