

Responsible Innovation Plattform

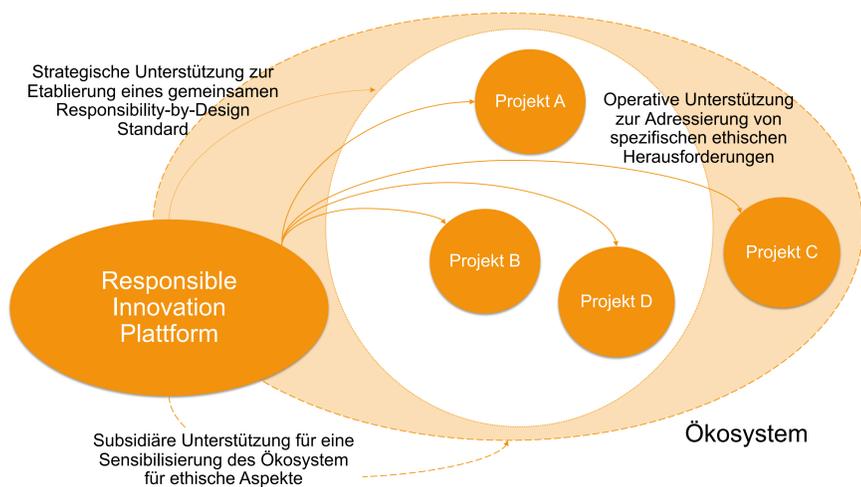
Verantwortungsvolles Innovieren für vertrauenswürdige KI

Sabrina Blank, M.A. & Prof. Dr-Ing. Christian Herzog, M.A. (Projektleitung)

Forschungsfelder der Responsible Innovation Plattform

Die Responsible Innovation Plattform steht für eine **frühzeitige, partizipative** und **iterative Analyse** ethischer, rechtlicher und sozialer Aspekte (ELSA) während der Entwicklung **vertrauenswürdiger KI**. Dafür wird synergetisch auf drei Ebenen geforscht:

- Strategische** Integration von ELSA mit dem Responsibility-by-Design Standard (CEN CWA 17796:2021),
- Operative** Adressierung von ELSA im Projekt,
- Subsidiäre** Bereitstellung von Hilfsmitteln für die eigenständige sozio-ethische Analyse durch Entwickler*innen.



Quelle: Forschungsfelder der Responsible Innovation Plattform (eigene Darstellung)

Forschung an der operativen Adressierung von ELSA: Bias/Fairness

Die Implementierung von ELSA auf operativer Ebene bedarf einer **projekt-spezifischen Ausgestaltung**, da jedes Technik-Entwicklungsprojekt spezifische sozio-ethische Implikationen und Entwicklungskontexte aufweisen kann.

Explorativer Forschungsansatz:

Am Beispiel des Themas „Bias“ werden im **Projekt VIKOOB** (AP360, PI: Prof. Dr. Philipp Rostalski) potenzielle Bias-Eintrittspunkte, deren Dokumentation und Evaluation explorativ erforscht. Diese Erkenntnisse sollen ermöglichen, dass der KI-Algorithmus keinen vorurteilbehafteten Output generiert. Damit kann diese Forschung zur Vertrauenswürdigkeit der Lösung beitragen, da eine bessere **Interpretierbarkeit der Ergebnisse** ermöglicht wird.

Forschung an der subsidiären Unterstützung: Sensibilisierung und ethische Selbstreflexion

Methoden für eine eigenständige initiale sozio-ethische Analyse durch die Entwicklungsteams müssen an deren Bedürfnisse und Anforderungen angepasst sein. Zur Bedarfserhebung haben wir verschiedene Maßnahmen durchgeführt:

- Workshop mit dem Konsortium zur Bedarfserhebung
- Analyse von bereits bestehenden Tools
- Umfrage zur Rolle von Ethik in Entwicklungsprozess aus der Perspektive von Entwickler*innen



Quelle: Schwerpunktthemen im Workshop mit dem Konsortium (eigene Darstellung)

Die Entwicklung eines wachsenden online **Repositoriums zu ELSA in KI-MED** soll die vielfältigen Wünsche für eine selbstständige ethische Analyse unterstützen.

Gemeinsame Publikation aus dem Konsortium:

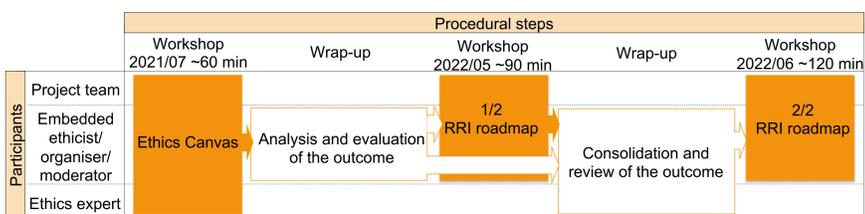
Petersen, E., Potdevin, Y., Mohammadi, E., Zidowitz, S., Breyer, S., Nowotka, D., Henn, S., Pechmann, L., Leucker, M., Rostalski, P., & Herzog, C. (2022). Responsible and Regulatory Conform Machine Learning for Medicine: A Survey of Challenges and Solutions. IEEE Access, 10, 58375–58418. doi.org/10.1109/ACCESS.2022.3178382



IEEEAccess-RRCMML

Forschung an der strategischen Integration von ELSA: Responsibility-by-Design Standard

Im **Projekt Assistierendes Bewegungstraining** (AP390, PI: Prof. Dr. Frank Steinicke) wurde der **Responsibility-by-Design Standard** angeleitet durch die **Responsible Innovation Plattform** adaptiert. Ethische Herausforderungen und Potenziale wurden in einer Reihe von Workshops erarbeitet, um Risiken für den Entwicklungserfolg sowie **entsprechende Handlungsoptionen frühzeitig zu identifizieren**. Die Ergebnisse resultieren in einer sogenannten **RRI Roadmap**, die ELSA als integralen Bestandteil des Entwicklungsprozesses plant.



Quelle: Prozessgestaltung der Adaption des Responsibility-by-Design Standard im Projekt (eigene Darstellung)

Die Auswertung des Prozesses zeigt, dass die Methodenwahl und Prozessgestaltung **effektiv** und **effizient** für die kollaborative sozio-ethische Kontextanalyse ist. Verbesserungspotenzial zeigt sich in der Zuordnung von Verantwortlichkeiten zur **Stärkung des Impacts**. Herausforderungen bei der Kontinuität sind im translationalen Kontext besonders prägnant.

Gemeinsam eingereichte Publikation:

Blank, S., Mason, C., Steinicke, F. & Herzog, C. (2023). Tailoring Responsible Research and Innovation to the Translational Context – The Case of AI-supported Exergaming for Assisted Movement Training. Submitted to Ethics and Information Technology (Springer Journal)

Ausblick: Auf dem Weg zum Vertrauenswürdigen KI-MED Ökosystem

Die Arbeit der Responsible Innovation Plattform versteht Vertrauenswürdigkeit begründet in rationalen Erwartungshaltungen als auch in der Übertragung von Motivationen und Werten — in einer epistemischen und moralischen Grundlage. Governancestrukturen, die diese beiden Grundlagen gleichermaßen unterstützen, sind daher ein essenziell für ein **vertrauenswürdigen KI-Med Ökosystem**.

Die Plattform hat ein wissenschaftliches Verstetigungskonzept mit Elementen eines „Ethics as a Service“ **Angebots zur projekt-individuellen, anwendungsnahen Berücksichtigung von ELSA** formuliert.

Eingereichte Publikation und Konzept in internationaler Kooperation:

Herzog, C., Blank, S., & Stahl, B. C. (2023). Trustworthy Medical AI Ecosystems—A Proposal for Ensuring Responsible Innovation Practices in AI-Based Medical Innovation. Submitted to Ethics and Information Technology (Springer Journal).