



INPUT-PAPIER

Herausforderungen und Chancen für KI-Ethik:

Von den Ergebnissen der letzten Roundtables zum definierten Handlungsfeld

Ein Input für den 3. Workshop der Roundtable-Reihe ethische KI-Entwicklung (RTeKI)

22. November 2022, Linda Schwarz, Julia Meisner (Gesellschaft für Informatik e.V.)

1 EINLEITUNG

Oft stehen in einem Innovationsprozess schnell mögliche Lösungen im Vordergrund. Beispielsweise hat unsere nicht repräsentative Umfrage ergeben, dass Softwareentwickler*innen vergleichsweise wenig für ethische Aspekte in ihrer Arbeit sensibilisiert sind. Zu diesem Problem kommen einem schnell mögliche Lösungen in den Sinn – beispielsweise Weiterbildungen, regelmäßige Meetings oder Leitlinien.

Im Design Thinking Prozess geht man davon aus, dass die zu eilige Suche nach Lösungen die eigene Kreativität einschränkt, da zu schnell und stark an Bekanntem festgehalten wird. Wie Henry Ford es formuliert haben soll: „If I had asked people what they wanted, they would have said faster horses.“ Im Design Thinking Prozess analysiert man ein Problem daher erst in mehreren Schritten aus verschiedenen Perspektiven, bevor explizit nach Lösungen gesucht wird. An diesem Prozess orientieren wir uns in der RTeKI-Reihe.

Im 3. Roundtable verweilen wir also noch einmal bei den **Herausforderungen und Chancen**, die wir in den letzten beiden Sitzungen zusammengetragen haben (*siehe Abschnitt 3*). Wir sammeln und verknüpfen diese in „How might we“-Fragen.

2 HOW-MIGHT-WE-FRAGEN

How-might-we-Fragen bilden die Grundlage und den Impuls, in diesem und den folgenden Roundtables auf (vorher vielleicht undenkbara Lösungs-)Ideen zu stoßen. Einerseits benennen How-might-we Fragen ein klares Handlungsfeld. Andererseits halten sie den Lösungsraum dabei für unkonventionelle Ideen offen und können bereits Ideen anregen. Dies wird im Folgenden erklärt.

How-might-we-Fragen sind in etwa wie folgt aufgebaut:

Wie schaffen wir es ...

... das Problem zu lösen?

... die Herausforderung zu bewältigen?

... das Bedürfnis zu erfüllen?



Die How-might-we-Fragen sollten möglichst fundiert und nuanciert formuliert werden – damit nehmen sie natürlich komplexere Formen an. Das allgemeine Problem, dass Softwareentwickler*innen zu wenig für ethische Aspekte in ihrer Arbeit sensibilisiert sind, kann zum Beispiel so formuliert werden:

*Wie schaffen wir es, dass Softwareentwickler*innen ihren Code und die Datenbasis selbstverständlich und eigenverantwortlich auf unbeabsichtigten Bias prüfen?*

Inspirierend wird es, wenn wir das formulierte Problem mit ähnlichen Situationen vergleichen – und wir damit an bekannte Situationen, Strukturen oder auch Gefühle anknüpfen. Beispielsweise kam im vorherigen Roundtable die Frage auf, was aus Erfahrungen mit Privacy-by-Design-Anforderungen gelernt werden könne. Die Fragen können aber auch in einen ganz neuen Kontext gesetzt werden.

Obige Frage kann dann zum Beispiel wie folgt aussehen:

*Wie schaffen wir es, dass Softwareentwickler*innen es als genauso selbstverständlich empfinden, ihren Code auf unbeabsichtigten Bias zu überprüfen, wie sich morgens ihren Kaffee einzuschmecken?*

Wir sind in diesem Beispiel auf der Alltagsebene der Softwareentwickler*innen angelangt – und haben damit eine präzisere Vorstellung, wo mögliche Lösungen ansetzen könnten.

3 BISHERIGE ERGEBNISSE: IDENTIFIZIERTE HERAUSFORDERUNGEN UND CHANCEN

Die folgende Übersicht soll dabei unterstützen, die How-might-we-Fragen zu formulieren. Sie ist eine Zusammenfassung der Ergebnisse der letzten beiden Roundtables. Die Probleme und Chancen stammen also sowohl aus Erfahrungen, die wir Teilnehmenden etwa in den 6-W-Fragen untereinander ausgetauscht haben, als auch von subjektiven Perspektiven, die wir im Persona-Building und in der Umfrage ergründet haben. Der Fokus liegt auf Problemen, die wir innerhalb der Roundtable-Reihe tatsächlich angehen können.

Herausforderungen

Unklares Ziel-Verständnis

Ethik ist per se uneindeutig. Sie kann nicht einfach gemessen werden und ist nicht „richtig“ oder „falsch“. Was z. B. Fairness bedeutet und welche Anforderungen damit einhergehen, hängt auch von der verwendeten KI und den Bereichen ab, in denen diese eingesetzt werden soll.

Viele Begriffe und Anforderungen sind unkonkret. Für Mitarbeitende ist es daher schwierig, oft verwendete Ethik-Begriffe (wie Transparenz, Fairness, Sicherheit, Vielfalt) umzusetzen.

Fehlendes Wissen

Ein persönlicher Bias wird oft nicht erkannt. Mitarbeitende sind sich eigenen Vorurteilen mitunter nicht bewusst oder ihnen fehlen Erfahrungen, um zum Beispiel mögliche Diskriminierungen zu erkennen.



Kaum jemand versteht, wie KI funktioniert. Damit bleibt auch unklar, auf welcher Grundlage Ergebnisse zustande kommen und inwiefern diese ethischen Prinzipien entsprechen.

Viele Mitarbeitende kennen die ethische Leitlinie des eigenen Unternehmens nicht, wenn vorhanden.

Mitarbeitenden ist unklar, wie ethische Richtlinien umgesetzt werden sollen. Es fehlt beispielsweise an praktischen Erläuterungen.

Fehlentscheidungen. Entscheider*innen haben mitunter zu wenig Expertise und praktische Einblicke, sind aber für wesentliche Entwicklungen verantwortlich.

Fehlende Relevanz

Der Begriff „Ethik“ ist bei vielen vorbelastet. Er wird z. B. als abstrakt oder als Marketing-Begriff wahrgenommen. Mitarbeitende können schwer motiviert werden, sich damit auseinanderzusetzen.

Mitarbeitende fühlen sich nicht verantwortlich, Ethik umzusetzen. Mitarbeitende verbinden mit Ethik v. a. Mehrarbeit, die mit ihrer eigenen Arbeit und ihren Aufgaben nicht zusammenhängt. (Z. B. Softwareentwickler*innen, die in ihrer Ausbildung nicht mit ethischen Fragen in Kontakt gekommen sind. Sie wollen „einfach nur programmieren“.)

Ethik in die KI-Entwicklung zu integrieren, kostet. Mitarbeitenden und dem Management ist unklar, inwiefern deren Berücksichtigung einen unternehmerischen Mehrwert bietet. Es besteht Zeit- und Gelddruck. Oft stehen (kurzfristige) finanzielle Interessen einer intensiven Berücksichtigung von KI-Ethik im Weg.

Kund*innen fordern keine KI-Ethik ein. Für sie ist sie kein funktionaler Bestandteil eines Produkts, die Berücksichtigung von KI-Ethik gilt (noch) nicht als Qualitätsmerkmal.

Komplizierte Umsetzung

KI-Ethik umzusetzen, ist komplex. Denn Entwicklungsprozesse verteilen sich oft auf viele Ebenen und Mitarbeitende und laufen über viele Monate und Jahre. (Zum Beispiel Softwareentwickler*innen, die keinen Einblick in die Datenlage haben und nur für eines von vielen Features eines Produkts verantwortlich ist.) Es ist schwierig, KI-Ethik dabei zielführend und effizient umzusetzen.

Unklare Verantwortung. Die Aufgabe, KI-Ethik umzusetzen, wird immer weiter delegiert – niemand fühlt sich dafür verantwortlich und setzt sie um.

Prozessmüdigkeit. Wird KI-Ethik mit intransparenten Vorgaben und ausartenden Diskussionen verbunden, fühlen sich Mitarbeitende bevormundet oder ermattet und daher wenig motiviert, sich mit dieser auseinanderzusetzen.

Unternehmen arbeiten oft global und in vielfältigen Abhängigkeiten. Es ist schwierig, diese verschiedenen Voraussetzungen konsequent bei der Umsetzung von KI-Ethik zu berücksichtigen.



Chancen

KI-Ethik ist ein relevanter Trend. Und wird als solcher auch schon von Kund*innen und Management erkannt.

Regulierungsmaßnahmen wie der AI Act werden kommen. Sie werden viele Unternehmen dazu verpflichten, sich mit KI-Ethik auseinanderzusetzen.

KI-Ethik ist ein marktwirtschaftlicher Vorteil. Er hilft Unternehmen dabei, sich von der nationalen und internationalen Konkurrenz abzuheben.

Jeder Mensch kann Ethik. Um Ethik ins Gespräch zu bringen, kann an persönliche Erfahrungen und Einstellungen angeknüpft werden. Statt Regeln können handlungsleitende Werte motivieren.

Unternehmen werden diverser. Sie bieten damit sehr gute Voraussetzungen für einen interdisziplinären und transkulturellen Austausch.

Aus bisherigen Prozessen lernen. Es gibt bereits Erfahrungen, z. B. Security by Design oder Privacy by Design umzusetzen.

Der technologische Fortschritt schreitet voran. Er bietet damit auch vielfältige technische Möglichkeiten, Werte umzusetzen.

Alle Aspekte kreisen um die Frage, wie wir Verantwortung von Mitarbeitenden wecken können, damit diese ethische Aspekte in der KI-Entwicklung reflektieren und umsetzen. Es ist eine lange Liste an Ergebnissen. Im 3. Roundtable wird es also auch darum gehen, abzuwägen und auszusortieren – und damit zu entscheiden, wofür wir in den kommenden Roundtables Lösungen entwickeln.