

AI FORA – The smart dystopia?

Elisabeth Späth, David Wurster

Gewährung von Asyl und Sozialdienstleistungen für Geflüchtete unter Einsatz von Künstlicher Intelligenz (KI)

In immer mehr nationalen Wohlfahrtsystemen werden staatliche Leistungen mit computertechnischer Unterstützung verteilt. Das AI FORA-Projekt (Laufzeit Mai 2021 – Dezember 2024) beschäftigt sich aus soziologischer Sicht mit Fairnessproblemen bei der Gewährung von Sozialdienstleistungen, unter Berücksichtigung des (möglichen) Einsatzes Künstlicher Intelligenz (KI).

KI-Systeme zeichnen sich durch ihre besonderen Fähigkeiten zur „statistischen Profilerstellung“ aus¹, und können genutzt werden, um z. B. die „berechtigten“ Förder-/Unterstützungsmaßnahmen für bestimmte Profile zu errechnen.

Voreingenommenheit innerhalb der algorithmischen Entscheidungsfindung

Gleichzeitig beruhen diese „Profile“ – und die daraus resultierende Differenzierung des Menschen – in der Regel nur auf probabilistischen Zahlen. Dies führt häufig zu Problemen wie Voreingenommenheit und Unfairness innerhalb der algorithmischen Entscheidungsfindung sowie letztendlich zu dem Problem der Verantwortlichkeit für mögliche Verzerrungen oder Diskriminierungen². Das partizipativ-angelaygte AI FORA-Projekt ist diesen Zusammenhängen in verschiedenen Fallstudienländern (u. a. Deutschland, Estland, Spanien, Indien und den USA) auf der Spur und lädt Betroffene ein, sich mit ihren Sichtweisen und Erfahrungen einzubringen. Der Ansatz partizipativer und inklusiver Technikgestaltung im Projekt wird mittels empirischer sozialwissenschaftlicher Forschung, Modellierung und Simulation umgesetzt. Ergebnisse werden mit Entscheidungsträger*innen aus Politik und Verwaltung diskutiert und evaluiert.

Die deutsche Fallstudie:

In der deutschen Fallstudie von AI FORA geht es um die Gewährung von Asyl und den Einsatz von KI in diesem Kontext sowie um Sozialdienstleistungen für

Geflüchtete. Anknüpfungspunkt sind hier diverse KI-Projekte des Bundesamtes für Migration und Flüchtlinge (BAMF). In der Digitalisierungsagenda 2022 beschreibt Antje Kiss, damalige Leitung der IT-Abteilung im BAMF, die zunehmende Relevanz von KI-Technologien: „Große Bedeutung im Asylbereich haben zukünftig Assistenzsysteme, die auf Künstlicher Intelligenz beruhen und das digitale Identitätsmanagement, das eine für alle Seiten bestmögliche und gerechte Asylantragsbearbeitung verwirklicht“ (BAMF 2021)³. Ein Beispiel ist die Nutzung der Dialekt-/Sprachanalysesoftware, im Einsatz seit 2017, mit deren Unterstützung die Herkunft von Asylwerbenden verifiziert werden soll⁴.

Empirische Forschung: Qualitative Interviews und interaktive Workshops

Ziel dieser Fallstudie ist es, mithilfe eines partizipativen Ansatzes, die bereits bestehenden Problemen in dem sehr komplexen wie dynamischen Feld der Asyl- und Migrationspolitik zu verstehen sowie die Einstellungen von Betroffenen bzgl. des (möglichen) KI-Einsatz aufzunehmen. Im Rahmen dieser Fallstudie des dreijährigen Forschungsprojekts wurden qualitative Interviews sowie insgesamt drei Workshops durchgeführt. Hierbei wurde keine bestimmte Zielgruppe in den Fokus genommen (hinsichtlich z. B.

³ Bundesamt für Migration und Flüchtlinge (2021). Digitalisierungsagenda 2022. https://www.bamf.de/SharedDocs/Anlagen/DE/Digitalisierung/broschuere-digitalisierungsagenda-2022.pdf?__blob=publicationFile&v=4 Download 15.06.2024

⁴ Bundesamt für Migration und Flüchtlinge (2018). Integriertes Identitätsmanagement – Plausibilisieren, Datenqualität und Sicherheitsaspekte. Einführung in das IDM-S Tool Auslesen von mobilen Datenträgern (AmD). Schulung AVS-Mitarbeiter. https://fragenstaat.de/dokumente/9650-schulung_avs_kurz/ Download 15.06.2024

¹ Fourcade, M. (2016): Ordinalization: Lewis A. Coser Memorial Award for Theoretical Agenda Setting 2014. *Sociological Theory* 34(3), S. 175-195.

² Cambridge: Harvard University Press.



Elahe Janavi: Cold Anger.

Geschlecht, Schutzstatus, Herkunftsland), um ein möglichst diverses und komplexes Bild der Erfahrungen sowie Einstellungen Geflüchteter gegenüber KI-Nutzung zu gewinnen. Aufgrund ethischer sowie rechtlicher Gründe wurden jedoch keine Minderjährigen sowie überwiegend Personen, deren Asylverfahren bereits abgeschlossen waren, befragt.

Zu Beginn des Forschungsprojekts wurden fünf explorative Sondierungsgespräche mit Mitarbeiter*innen aus Wohlfahrtsverbänden sowie Flüchtlingsbeauf-

tragten der Kirche geführt, um ihre Perspektiven auf politische, soziale sowie technische Entwicklungen im Hinblick der Migrations- und Asylpolitik in den letzten 5-10 Jahren zu erfahren. Zudem wurden sieben Interviews mit Geflüchteten (aus Afghanistan, Syrien, Iran) sowie Unterstützer*innen geführt, in denen es um ihre Erfahrungen in den einzelnen Prozessen des Asylverfahrens ging, den erlebten Zugang zu Integrations- und Sozialdienstleistungen, sowie ihre Ansichten zu Verbesserungen im System (technische, soziale, organisatorische Aspekte).

Neben Literaturrecherchen (z. B. Stand des KI-Einsatz in Asylverfahren und Integrationsprozessen im In- und Ausland; Digitalisierung der öffentlichen Verwaltung in Deutschland), dienten diese Interviews dazu, die Schwerpunkte der interaktiven Workshops zu setzen.

Workshops mit Geflüchteten, Funktionär*innen und politisch Verantwortlichen

Zum ersten Workshop, der im Juni 2022 stattgefunden hat, waren Geflüchtete (aus

Afghanistan, Syrien, Iran) und (überwiegend) ehrenamtliche Unterstützer*innen, u. a. Ämterlots*innen, eingeladen, um ihre Erfahrungswelt und Einstellungen zu Fairness-Aspekten und Nutzung von (KI-) Technologie zur Unterstützung administrativer Prozesse kennenzulernen. Im Rahmen von zwei Fokusgruppen wurden sie zu ihren a) Erfahrungen in den Behörden (wie BAMF, Ausländerbehörde, Jobcenter) sowie deren Zusammenarbeit, b) zur Rolle/Bedeutung der „deutschen“ Kultur oder Werte in diesen Prozessen (z. B. Bürokratie, Sicherheitsbestreben) sowie c) den erfahrenen/erlebten Bewertungskriterien und ihre Wünsche bezüglich Veränderungen in Asylverfahrens- und Integrationsprozessen befragt. Zwei wichtige Aspekte, die genannt sowie im Rahmen eines Worldcafés am darauffolgenden Tag diskutiert wurden, waren zum einen die individuellen sowie kollektiven Möglichkeiten sich zu beteiligen, zum anderen, die wahrgenommene bzw. erlebte Ungleichbehandlung/-bewertung von ukrainischen und nicht-ukrainischen Geflüchteten. Diese Aspekte sind für den thematischen Fokus der Fallstudie/des Forschungsprojekts – Bewertungspraktiken und „Kategorisierungen“ von Menschen, denen aufgrund bestimmter Kriterien bestimmte Rechte eingeräumt werden – relevant.

Im Juli 2023 waren hauptamtlich tätige Expert*innen, insbesondere Vertreter*innen aus Wohlfahrtsverbänden, Anti-Diskriminierungsverbänden und Beauftragte für Integration und asylrechtliche Fragen, zu einem Workshop eingeladen. Schwerpunkt dieses Formats war, zum einen, Entscheidungsprozesse in den einzelnen Schritten (von Ankunft in Deutschland, Asylantrag, Anhörung, Entscheidung etc.) näher zu beleuchten und, zum anderen, die verschiedenen Herausforderungen und Barrieren sowie Möglichkeiten und Unterstützung für Geflüchtete zu verstehen. Im Anschluss wurde diskutiert, an welchen Stellen KI-Unterstützung hilfreich wäre bzw. an welchen Stellen KI nicht zum Einsatz kommen sollte.

Zum Abschluss fand im Juli diesen Jahres ein Workshop mit politischen Entscheidungsträger*innen v. a. auf Landesebene aus den Themengebieten KI, Digitalisierung und Migrationspolitik sowie mit Personen aus der Praxis, insbesondere aus dem Schnittstellenbereich von Integration und Qualifizierung/Arbeit, statt. Hauptfokus dieses Work-

shops bestand in der Vorstellung der Projektergebnisse sowie in der gemeinsamen Evaluierung dieser im Rahmen von interaktiven Gruppendiskussionen. Die Ergebnisse des Workshops wurden im Rahmen einer Podiumsdiskussion, als Teil der Veranstaltung „European Workshop on Algorithmic Fairness“ (<https://2024.ewaf.org/>) vorgestellt bzw. zusammen mit dem Publikum – Wissenschaftler*innen aus unterschiedlichen Disziplinen (Informatik, Jura, Philosophie, Sozialwissenschaften) – diskutiert.

Wichtige Ergebnisse:

1. Bestehende (Fairness-)Problematiken, die sich auf Geflüchtete und deren „Lebensweg“ in Deutschland klar nachteilig auswirken sind Diskriminierungs- und Rassismuserfahrungen sowie mangelnde Digitalisierung bzw. überbordende Bürokratie. Verstärkt werden diese durch die diskrepante (Nicht-)Nutzung von (KI)-Technologien: a) Während das BAMF „High-Tech“-Projekte initiiert und implementiert, gibt es innerhalb der Strukturen wichtiger Behörden, wie Ausländerbehörden, wenig digitalisierte Prozesse. b) Die normative bzw. politische Dimension von Bewertungskriterien (z. B. Einteilung in (un-)sicheres Herkunftsland; Unterschiede bzgl. Herkunftsländer, s. Ukraine) kann durch den KI-Einsatz verstärkt werden, durch den Fokus auf bestimmte Bevölkerungsgruppen oder ein bestimmtes Merkmal – Herkunftsland/-region. Dies kann dazu führen, dass andere Kriterien, im Sinne der Einzelfallentscheidung, nicht im vollen Umfang mitberücksichtigt werden.
2. Während Werte, wie Fairness, Datenqualität, Effizienz, aber auch Transparenz und Vertrauen, für alle Akteure eine wichtige Rolle spielen, stehen die Akteure unterschiedlich zu diesen – aufgrund von Arbeitsschwerpunkten oder unterschiedlichen Interessen (z. B. BMI, BMAS, Migrant*innenorganisationen). Diese Werte, z. B. Fairness und Effizienz, können, müssen aber nicht im Konflikt stehen. Jedoch bewirkt „Silo-Denken“, dass einseitig vorteilhafte (technologische), oft nur kurzfristige, Lösungen gefunden werden.
3. Unterschiedliche Perspektiven von Geflüchteten, Expert*innen, Unterstützer*innen, sollten zusam-

mengeführt werden, um KI-Potentiale voranzubringen. Beispielsweise hatten Geflüchtete eine überwiegend positive oder optimistische Einstellung gegenüber KI geäußert, da sie zu mehr Objektivität und Effizienz beitragen könne. Ein genanntes konkretes Beispiel ist der Abgleich bzw. Einstufungsgrad ausländischer Studiengänge und Berufsausbildungen für deren Anerkennung.

Die Ergebnisse der empirischen Forschung fließen zurzeit in die Arbeiten zur Modellierung und Simulation von Szenarien der Technikgestaltung für (eine) „bessere KI“ ein, welche Perspektiven der Betroffenen abbilden⁵. Die erste Publikation zu dieser Fallstudie, sowie alle anderen Fallstudien, finden Sie demnächst in dem vom Springer-Verlag als Open-Access veröffentlichtem Buch „Participatory Artificial Intelligence in Public Services. From Bias to Fairness in Assessing Beneficiaries“⁶. Bei Fragen oder Rückmeldungen können Sie sich bei Projektmitarbeiterin Elisabeth Späth (espaeth@uni-mainz.de) melden.

Elisabeth Späth, David Wurster sind Mitarbeitende im AI FORA-Projekt, das im Rahmen der von der VolkswagenStiftung geförderten Initiative „Künstliche Intelligenz – Ihre Auswirkungen auf die Gesellschaft von morgen“ an der Johannes Gutenberg Universität Mainz, Institut für Soziologie, Lehrstuhl für Technik- und Innovationssoziologie, Soziale Simulationen, durchgeführt wird. <https://www.ai-fora.de/>

⁵ Ahrweiler, P. (2017): Simulationsexperimente realexperimenteller Politik – der Gewinn der Zukunftsdimension im Computerlabor. In: Boesch, S., Gross, M. and W. Krohn (eds.): Experimentelle Gesellschaft. Baden-Baden, 199-237.

⁶ Späth, E. The AI use in the asylum procedure in Germany: exploring perspectives with refugees and supporters on assessment criteria and beyond. In: Ahrweiler, P. (Hrsg.), Participatory Artificial Intelligence in Public Services. From Bias to Fairness in Assessing Beneficiaries, Artificial Intelligence, Simulation and Society Series, Cham: Springer.



Elahe Janavi: Rekless.