

# Können Algorithmen Helfer sein?

## Künstliche Intelligenz im Einsatz für eine verantwortungsvolle Ressourcenallokation

CHRISTOF WOLF-BRENNER, KNOW-CENTER, SEBASTIAN DENNERLEIN,  
UNIVERSITY OF TWENTE, ROBERT GUTOUNIG, FH JOANNEUM, STEFAN  
SCHWEIGER, BONGFISH, VIKTORIA PAMMER-SCHINDLER, TU GRAZ – VOM  
27.10.2021

Wie können knappe Güter zum Wohle möglichst vieler am besten eingesetzt werden? Diese Frage führt zum Allokationsproblem. Es ist so alt wie die Menschheit selbst. Zwei Beispiele dafür werden in Österreich öffentlich diskutiert: Zum einen die automatisierte Einteilung von Arbeitssuchenden als Basis für die Freigabe von Förderungen, auch bekannt als "AMS-Algorithmus". Zum anderen die Einschätzung der Überlebenschancen von Patient\*innen mit lebensbedrohlichen Verläufen von Covid-19 zur Zuteilung von Intensivbetten -"Triage" genannt. In beiden Fällen unterstützen Algorithmen Expert\*innen, damit diese fundierte Entscheidungen zur Verteilung von Ressourcen treffen.

Auch wenn die Beispiele fundamental unterschiedlich anmuten, eint sie, dass dafür jeweils mittels Beispieldaten Modelle erstellt wurden. Füttert man diese Modelle mit neuen, gleichartigen Daten, erbringen sie ein den Beispieldaten entsprechendes Ergebnis. Im Falle des AMS-Algorithmus eine grobe Kategorisierung in Gruppen mit hoher, mittlerer oder niedriger Wahrscheinlichkeit, in den nächsten Monaten wieder eine Anstellung zu finden. Bei der Triage werden Patient\*innen anhand klinischer Parameter in 21 Gruppen eingeteilt, die unterschiedliche Wahrscheinlichkeiten repräsentieren, im Krankenhaus an Covid-19 zu sterben.

Die Ergebnisse werden jeweils von den Expert\*innen genutzt, um Entscheidungen mit tiefgreifenden Konsequenzen zu fällen, nämlich: "Wird der Arbeitssuchende mit Förderungen unterstützt?" Und: "Bekommt die Patientin ein Intensivbett?"

Der AMS-Algorithmus steht seit 2018 in der Kritik. Bei der Triage ist sie weitgehend ausgeblieben. So liegt die Vermutung nahe, dass einer der Gründe dafür in der im Detail unterschiedlichen Ausgestaltung der Algorithmen zur Unterstützung von recht ähnlichen Entscheidungen liegt.

AMS-Algorithmus und Triage im Vergleich Um zu ergründen, welche moralischen Problembereiche den beiden Beispielen innewohnen, ist es hilfreich, eine systematische Untersuchung durchzuführen. Dazu wurde von Grazer Forscher\*innen das soziotechnische Reflexionsframework SREP entwickelt. Es zielt darauf ab, zum Design von verantwortungsvollen KI-Anwendungen anzuleiten. Dabei werden relevante Interaktionen von Mensch und Maschine während Entwicklung und Verwendung von algorithmisch unterstützten Systemen vor dem Hintergrund ethischer Grundbausteine reflektiert und potenzielle Probleme festgestellt. Beispielhaft betrachten wir hier die Prinzipien "Transparenz" und "Fairness", um die Auswirkungen unterschiedlicher Designentscheidungen für die Klassifizierungen bei Triage und AMS-Algorithmus zu beleuchten.

Im Kontext von KI heißt Transparenz, dass die innere Funktionsweise eines Algorithmus, der aus dem Input einen Output generiert, für Menschen verständlich und nachvollziehbar ist. Die generelle Vorgehensweise beim AMS-Algorithmus ist zwar, verständlich dokumentiert, veröffentlicht, über die meisten der angewandten Klassifizierungsmodelle ist jedoch wenig oder gar nichts bekannt. Bekommen also AMS-Berater\*in oder Arbeitssuchende nach Preisgabe ihrer personenbezogenen Input-Daten das errechnete Ergebnis präsentiert, ist nach aktuellem Kenntnisstand für beide nicht nachvollziehbar, warum es so ausfiel, wie es ausfiel. Dies verhindert, mit dem Output des Algorithmus sinnvoll weiterarbeiten zu können.

Ganz anders bei der Auswahl von Patient\*innen in der Zuteilung von Intensivbetten. Die Patient\*innen werden nach der klinischen Erfolgsaussicht einer Intensivtherapie priorisiert. Das Fachpersonal ist angehalten, eine solche Einschätzung mittels nachvollziehbarer Kriterien vorzunehmen und dabei beispielsweise auf den 4C-Mortalitäts-Score zurückzugreifen. Transparenz wird dabei als Maßstab vorausgesetzt. Unterschiedliche Ergebnisse sind eindeutig auf Eingabeparameter rückführbar und ermöglichen so eine vergleichbare Bewertung der Erfolgsaussichten von Patient\*innen einer Intensivtherapie.

Sortiert der AMS-Algorithmus Antragstellende in die Gruppe der Förderungswürdigen ein, liegt die Chance, dass diese Personen innerhalb von etwa einem halben Jahr eine Anstellung finden, per Definition zwischen 25 und 66 Prozent. Was aber wäre, wenn eine Einschätzung knapp darüber oder darunter läge? Die betroffenen Personen würden entweder wenig oder gar nicht gefördert werden. Zwei Menschen, die beinahe idente Jobchancen haben, müssen dank unterschiedlicher Klassifizierung mit stark

ungleicher Behandlung rechnen, weil der AMS-Algorithmus sie nicht als förderungswürdig erfasst hat. Sofern kein Mensch korrigierend eingreift, fällt er die Entscheidung, ob eine Person gefördert wird oder nicht.

Auch beim Mortalitäts-Scoring der Triage wird klassifiziert. Im Gegensatz zum AMS-Algorithmus gibt es beim 4C-Risikomodell jedoch 21 unterschiedliche Klassen, denen jeweils eine Mortalitätswahrscheinlichkeit zugeordnet ist. Die Differenz zwischen zwei Risikoklassen beläuft sich durchschnittlich auf etwa 4,5 Prozent. Abgesehen davon gibt es zwischen den beiden Algorithmen keine fundamentalen Unterschiede in der generellen Funktionsweise.

Der Vorsprung an Fairness bei der Triage-Unterstützung ergibt sich aus der Anwendung der Ergebnisse aus dem Algorithmus. Während beim AMS pro Person eine Klasse prognostiziert wird und darauf basierend eine Investitionsentscheidung erfolgt, wird bei der Triage in der Regel paarweise verglichen: Welcher von zwei Patient\*innen hat den niedrigeren Mortalitäts-Score, somit die besseren Überlebenschancen und bekommt daher das Intensivbett?

Die verfügbaren Ressourcen werden optimal genutzt, da auf Basis der klinischen Einschätzung von Expert\*innen, unterstützt durch Scoring-Modelle, mittels Priorisierung entschieden wird, wer auf die Intensivstation verlegt wird. Beim AMS hingegen gibt es keine Priorisierung: Auch der aussichtsreichste Kandidat unter Menschen mit niedriger Einstellungschance bleibt für den Algorithmus förderungswürdig. Die Antragstellenden müssen sich auf die wachsamen Augen der Berater\*innen beim AMS verlassen, damit diese korrigieren, was die Rationalisierungsalgorithmen nicht leisten können.

Erfahrungsaustausch für ethisch vertretbare Lösungen Dieser Beitrag lobt weder die Algorithmen zur Sterblichkeitsschätzung, noch polemisiert er gegen den AMS-Algorithmus. Es ist der Versuch zu zeigen, dass nicht nur die Algorithmen selbst, sondern auch die Art und Weise, wie ihre Ergebnisse kommuniziert und angewandt werden, die Qualität von Entscheidungen zur Lösung von Allokationsproblemen maßgeblich beeinflussen. Um zu klären, in welchen Aspekten unterschiedliche Lösungsansätze abweichen, sollte auf Basis geeigneter Frameworks systematisch bedacht werden, was an genau jenen Stellen passiert, an denen Mensch und Maschine interagieren. Steht dieser Interaktionsprozess im Einklang mit unserem Verständnis von verantwortungsvollem und ethisch vertretbarem KI-Einsatz?

Wir verstehen diesen Beitrag als Plädoyer für einen Erfahrungsaustausch zwischen Entwickler\*innen und Nutzer\*innen von KI-basierten Systemen. Es ist unerlässlich, in Österreich, in Europa aber auch weltweit eine Gemeinschaft aufzubauen, die Erfahrungen sammelt, gemeinsam ethische Fragestellungen reflektiert und mögliche Lösungsansätze dokumentiert. Damit kann sie bei der Entwicklung und Nutzung von KI-Tools mit Rat und Tat zur Seite stehen. Dabei sollen die gesammelten Daten offengelegt und der Allgemeinheit zur Verfügung gestellt werden. So können jene Entwickler\*innen, die an ähnlichen Lösungen arbeiten, von vergangenen Fehlern lernen.

Im Lichte jüngster Entwicklungen seitens der EU zur Regulierung von KI-Anwendungen wird es nicht mehr genügen, sich nachträglich mit soziotechnischen Fragen auseinanderzusetzen. Vielmehr ist es nötig, auf Erfahrungen mit vergleichbaren Projekten zurückzugreifen, um tatsächlich behaupten zu können, bei KI-Entwicklungen nach bestem Wissen und Gewissen gehandelt zu haben.