Verantwortungsvolle Empfehlungssysteme für die medizinische Diagnostik

Daniel Schlör, Andreas Hotho

1. Einleitung

Die frühe Entwicklung und Einführung von Krankenhaus- und Gesundheitsinformationssystemen hat die fortschreitende Digitalisierung von Prozessen in Krankenhäusern gefördert. Viele dieser Prozesse, die früher Schriftverkehr und telefonische Absprachen erforderten, sind heute in IT-Lösungen integriert und erfordern die Interaktion von Ärzt*innen und medizinischem Personal mit entsprechenden Schnittstellen und Tools. Diese Umstellung auf digitales Datenmanagement und Prozessunterstützung kommt der Versorgung der Patient*innen zwar in vielerlei Hinsicht zugute, erfordert aber von den Ärzt*innen eine genaue digitale Erfassung aller relevanten Informationen für Abrechnungs- und Dokumentationszwecke und damit oft eine Auswahl aus einer unüberschaubar großen Menge an Möglichkeiten. Das beansprucht viel Zeit, die sonst für die ärztliche Versorgung der Patient*innen aufgewendet werden könnte. Die systematische Erfassung von Gesundheitsdaten über einen langen Zeitraum hinweg bietet jedoch Möglichkeiten, diesen Prozess zu verbessern und das medizinische Personal durch die Einführung von Empfehlungssystemen zu unterstützen.

Betrachten wir zum Beispiel eine Ärztin in einem Krankenhaus, die eine Untersuchung bei einem Patienten durch eine andere Abteilung durchführen lassen möchte. Dazu muss sie eine entsprechende Anfrage stellen, die die Anamnese und die medizinische Fragestellung enthält, und die richtige Untersuchung auswählen. Diese Auswahl aus einer großen Menge von Untersuchungen kann durch ein Empfehlungssystem unterstützt werden, das entweder inhaltsbasiert arbeitet, oder auf ähnlichen Fällen aus der Vergangenheit basiert. Inhaltsbasierte Empfehlungen könnten in diesem Beispiel etwa

Anamnese, medizinische Fragestellung, mögliche medizinische Vorgeschichte des Patienten sowie anderes medizinisches Wissen berücksichtigen, um eine möglichst passende Empfehlung für die Untersuchung auszusprechen. Für Empfehlungen anhand ähnlicher Fälle würden beispielsweise hinsichtlich der vorherigen Behandlungssequenz oder Patientenzustand ähnliche Fälle betrachtet werden und die in diesen Fällen durchgeführten Untersuchungen empfohlen werden.

Beide Ansätze haben technische Vor- und Nachteile, die in Abschnitt 2.1 genauer benannt werden, sind aber auch für unterschiedliche Anwendungszwecke verschieden gut geeignet. Betrachtet man beispielsweise speziell radiologische Untersuchungen, was genau den oben genannten Fall zwischen anfordernder und erfüllender Abteilung darstellt, zeigen sich große Unterschiede, je nach anfordernder Stelle: Während in der Onkologie beispielsweise häufig ähnlich gelagerte Fälle betrachtet werden, die dementsprechend auch gut durch Empfehlungen basierend auf ähnlichen Fällen unterstützt werden können, finden sich in der Notaufnahme oder Unfallchirurgie sehr unterschiedlich gelagerte Fälle, die eine einzelne Betrachtung basierend auf der konkreten Anamnese erfordern, insbesondere auch, weil häufig zunächst keine Krankengeschichte der Patient*in verfügbar ist.

Mit diesen verschiedenen Perspektiven eröffnen sich aber auch unterschiedliche ethische Fragen für den Einsatz von Empfehlungssystemen in diesem sensiblen medizinischen Kontext, die berücksichtigt werden müssen, um ein verantwortungsvolles Verhalten des Systems selbst sowie in der Interaktion mit dem medizinischen Personal sowie den Patient*innen zu gewährleisten. Diese Überlegungen betreffen beispielsweise das Krankenhauspersonal, etwa wenn es um deren Effizienz, mögliche Kosten- und Zeitersparnisse und den Konsequenzen für die Arbeitsplätze geht, aber auch die Patient*innen, die davon profitieren könnten. Fragestellungen hinsichtlich der Datenqualität und möglichen beispielsweise diskriminierenden Verzerrungen, die ein solches System verstärken könnte, aber auch Überlegungen, wer im Kontext eines solchen sozio-technologischen Systems welche Verantwortungen und damit Rechenschaftspflichten hat, sind hier relevant.

Ausgehend von diesen Überlegungen werden in diesem Beitrag Kriterien für ein verantwortungsbewusstes Empfehlungssystem im medizinischen Kontext aus einer anwendungsorientierten Perspektive skizziert und mögliche Designentscheidungen mit besonderem Fokus auf die bereits genannten Fragen sowie Sicherheit, Transparenz und Konformität hinsichtlich der anzuwendenden Regularien betrachtet. Hierfür nehmen wir aus der informa-

tischen Perspektive den Standpunkt für ein konkretes Anwendungsszenario ein und werfen dafür relevante Fragestellungen auf, die von den technischen Aspekten losgelöst aus ethischer Sicht betrachtet werden können.

2. Grundlagen

In diesem Kapitel sollen zunächst die Grundlagen zu Empfehlungssystemen sowie die hier betrachtete Operationalisierung ethischer Aspekte vorgestellt werden.

2.1 Empfehlungssysteme

Empfehlungssysteme sind computergestützte Algorithmen und Systeme, die Anwender*innen auf Grundlage ihrer Vorlieben, Interessen oder ihrem bisherigen Verhalten Artikel oder andere Objekte vorschlagen (Bobadilla et al. 2013). Diese stellen im klassischen Sinne beispielsweise Produkte in Online-Shops (Fischer, Zoller und Hotho 2021), Bücher (Mooney und Roy 2000), Musik (Song, Dixon und Pearce 2012) und Videos (Gomez-Uribe und Hunt 2015), aber auch weniger typische Objekte, wie beispielsweise Schlagworte (Jäschke et al. 2007), Gegenstände in Computerspielen (Dallmann et al. 2021), medizinische Behandlungen (Sun et al. 2016) oder in unserem Beispiel medizinische Untersuchungen dar und sollen im Folgenden unter dem Begriff *Artikel* subsumiert werden. Empfehlungssysteme spielen eine entscheidende Rolle bei der Unterstützung der Anwender*innen z.B. bei der Navigation durch die großen Mengen an verfügbaren Informationen und verbessern so den Entscheidungs- oder Auswahlprozess (Zhang et al. 2019).

Empfehlungssysteme nutzen dafür Algorithmen und Techniken, um Daten der Nutzenden zu analysieren und häufig personalisierte Empfehlungen zu generieren. Durch die Berücksichtigung von Benutzungsfeedback (Shi, Larson und Hanjalic 2014) oder vorheriger Interaktionen (Adomavicius und Tuzhilin 2005) zielen solche Systeme darauf ab, den Anwender*innen relevante und hilfreiche Vorschläge zu unterbreiten. Der Hauptzweck von Empfehlungssystemen besteht darin, bei der Entdeckung neuer Artikel und Ressourcen unter anderem bezüglich der persönlichen Präferenzen zu unterstützen (Zhang et al. 2019), indem sie die sich teilweise entgegenstehenden Faktoren wie Genauigkeit, Neuartigkeit, Streuung und Stabilität der Empfehlungen balancieren (Bobadilla et al. 2013).

Hinsichtlich der verwendeten Methoden lassen sich Empfehlungssysteme in drei größere Kategorien einteilen: Systeme, die auf Collaborative Filtering (Adomavicius und Tuzhilin 2005) basieren, solche, die Content-based Filtering (Pazzani und Billsus 2007) nutzen, sowie Hybride Ansätze (Balabanović und Shoham 1997).

2.1.1 Collaborative Filtering

Collaborative Filtering ist eine weit verbreitete Technik in Empfehlungssystemen (Su und Khoshgoftaar 2009). Sie empfiehlt Artikel auf Grundlage der Vorlieben ähnlicher Benutzer*innen. Durch die Analyse von Verhaltensmustern und Ähnlichkeiten in den Interaktionen zwischen Benutzer*innen und Artikeln bzw. in den entsprechenden Bewertungen schlagen Collaborative Filtering Ansätze Artikel vor, die Benutzer*innen mit ähnlichen Vorlieben gefallen haben. Collaborative Filtering kann in zwei Hauptansätze unterteilt werden: nutzungsbezogenes und artikelbezogenes Collaborative Filtering (Aggarwal et al. 2016).

Nutzungsbezogenes Collaborative Filtering vergleicht die Vorlieben einer Zielnutzer*in mit denen ähnlicher Nutzer*innen, um Empfehlungen auszusprechen. Bei artikelbasiertem Collaborative Filtering hingegen werden ähnliche Artikel identifiziert und solche empfohlen, die denen ähneln, die der Zielbenutzer*in gefallen haben bzw. mit denen die Person zuvor interagiert hat.

Die Collaborative Filtering Ansätze haben sich zwar bei der Erstellung von personalisierten Empfehlungen als effektiv erwiesen, weisen aber auch Schwächen auf, wie zum Beispiel dem Kaltstartproblem (Schein et al. 2002), d.h. Schwierigkeiten bei der Empfehlung neuer oder inaktiver Benutzer*innen, und der Gefahr der Bildung von Filterblasen (Nguyen et al. 2014), d.h. Einschränkung der Empfehlungen auf eine begrenzte Auswahl. Auch sind für jede Nutzer*in in der Regel nur für wenige Artikel die Präferenzen bekannt, sodass man sehr viele Nutzende benötigt, um passende Empfehlungen aussprechen zu können.

2.1.2 Content-based Filtering

Bei Content-based Filtering werden Elemente auf der Grundlage ihrer Attribute oder Merkmale empfohlen. Solche Systeme analysieren den Inhalt oder die Merkmale bewerteter Artikel und gleichen sie mit den Präferenzen der Benutzer*in hinsichtlich anderer Artikel mit ähnlichen Merkmalen ab (Pazzani 1999).

Ein Vorteil inhaltsbasierter Methoden ist, dass Empfehlungen für neue Artikel gegeben werden können, auch wenn für diese noch nicht ausreichend Nutzungsinteraktionen vorhanden sind. Indem bewertete Artikel mit ähnlichen Attributen zur Bestimmung der Präferenz der nutzenden Person herangezogen werden, kann Content-based Filtering dennoch relevante Empfehlungen vorschlagen (Aggarwal et al. 2016).

Schwächen des Ansatzes sind nach Aggarwal et al. (2016) etwa die Abhängigkeit von Artikelmerkmalen wie Schlüsselwörtern oder anderem Inhalt, die zu offensichtlichen Empfehlungen und damit zu einer Verringerung der Vielfalt der empfohlenen Artikel führen kann. Außerdem setzt der Ansatz voraus, auf vorherige Bewertungen bzw. Interaktionen der gleichen Nutzer*in zugreifen zu können.

2.1.3 Hybride Ansätze

Hybride Empfehlungssysteme zeichnen sich durch die Kombination von Collaborative und Content-based Filtering aus. Diese Systeme nutzen die Vorteile beider Methoden, um genauere und vielfältigere Empfehlungen zu liefern (Burke 2002). Durch die Kombination von Nutzungspräferenzen (Collaborative Filtering) und Benutzer- und Artikelmerkmalen (Content-based Filtering) zielen hybride Ansätze darauf ab, die Empfehlungsqualität weiter zu verbessern, indem sie z. B. fehlende Bewertungen für das Collaborative Filtering einzelner Nutzer*innen auf Basis des Inhaltes vorhersagen und so die Einschränkungen der einzelnen Techniken vermeiden (Balabanović und Shoham 1997).

2.2 Betrachtung ethischer Aspekte

In den letzten Jahren hat die wachsende Bedeutung datengesteuerter Technologien auf der Grundlage des maschinellen Lernens Bedenken hinsichtlich ihrer Fairness, Transparenz und möglichen Voreingenommenheit aufgeworfen (Mehrabi et al. 2021):

Maschinelles Lernen, wie es für Empfehlungssysteme, aber auch im breiteren Kontext der Künstlichen Intelligenz (KI) verwendet wird, dient dazu, in Daten Muster zu bestimmen oder aus Daten Modelle zu lernen und diese für Vorhersagen oder Empfehlungen zu nutzen. Wenn die Trainingsdaten jedoch verzerrt sind, werden diese Verzerrungen auch vom Modell übernommen und spiegeln sich in dessen Ergebnissen wider (Chen et al. 2023).

Dies hat dazu geführt, dass sich eine wachsende Forschungsgemeinschaft mit den ethischen Auswirkungen von Systemen des maschinellen Lernens be-

schäftigt. In diesem Zusammenhang bietet die ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT) (Fox et al. 2023) als interdisziplinäre Konferenz eine Plattform für Wissenschaftler*innen aus verschiedenen Bereichen wie Informatik, Recht, Sozial- und Geisteswissenschaften, um sich mit den ethischen Herausforderungen der KI als sozio-technologisches System auseinanderzusetzen, und hat die in ihrem Titel formulierten Aspekte von Fairness, Verantwortlichkeit und Transparenz als mögliche Teilperspektiven verantwortungsvoller KI benannt.

Im Zusammenhang mit Empfehlungssystemen zielen Workshops und Tutorials wie FAccTRec (Fairness, Accountability, and Transparency in Recommender Systems) (FAccTRec 2022) oder Fairness and discrimination in recommendation and retrieval (Ekstrand, Burke und Diaz 2019) darauf ab, speziell diese Fragestellungen für Empfehlungssysteme zu erörtern, und folgen damit dem Vorbild von ACM FAccT, um die jeweiligen Aspekte speziell für den Anwendungsfall verantwortungsvoller Empfehlungssysteme zu untersuchen. Dabei werden die im Titel genannten Konzepte der Fairness, der Rechenschaftspflicht und der Transparenz um weitere wichtige Themen wie Verantwortung, Sicherheit, Compliance und die Auswirkungen fairnessbewusster und verantwortungsvoller Empfehlungen in Industrie und Forschung erweitert (FAccTRec 2022).

3. Responsible Recommender Design

In diesem Abschnitt stellen wir ein medizinisches Empfehlungssystem aus praktischer Sicht vor und erörtern die ethischen und rechtlichen Implikationen des allgemeinen Umfelds und der Designentscheidungen für das Empfehlungssystem.

3.1 Recommender Use Case

In Krankenhäusern gibt es mehrere Abteilungen, die auf bestimmte Diagnosen spezialisiert sind und anderen Abteilungen Dienstleistungen anbieten, beispielsweise Radiologie oder Labore. Um eine Diagnose für eine Patient*in zu erstellen, stellt die federführende Abteilung des Falles eine Anfrage an eine andere Abteilung, mit der sie eine bestimmte Untersuchung anfordert. Diese Untersuchung wird dann in der Abteilung durchgeführt, die die diagnostische Leistung anbietet, und die Ergebnisse werden an die anfordernde Abteilung

zurückgemeldet. Während in der Vergangenheit diese Diagnostikanfragen überwiegend telefonisch organisiert wurden, haben heute vor allem größere Krankenhäuser diesen Prozess mit IT-Lösungen digitalisiert. Am in diesem Beispiel betrachteten Klinikum können Ärzt*innen auf diese Weise Diagnosen aus der Radiologie anfordern. Für einen effizienten Anforderungsprozess hat sich das Klinikum entschieden, nur eine grobe Diagnosekategorie wie z.B. Röntgen des Handgelenks definieren zu lassen, anstatt das gesamte Spektrum der rund 2000 möglichen Untersuchungen darzustellen, die durchgeführt und abgerechnet werden können. Für die Dokumentation und Abrechnung muss jedoch das genaue diagnostische Verfahren erfasst werden. Deshalb beschäftigt das Klinikum medizinische Fachangestellte, die die Anfrage verfeinern, bevor sie diese an die entsprechende Abteilung weiterleiten. Dieser Ansatz hat zwei Nachteile, die durch Empfehlungssysteme adressiert werden können. Erstens nimmt die Verfeinerung durch das medizinische Personal Zeit in Anspruch, was bei dringenden Fällen im Gegensatz zu einem direkten Kontakt zwischen anfordernder und leistender Abteilung ein Problem darstellen kann, insbesondere wenn es Rückfragen an die anfordernde Abteilung gibt. Zweitens wird die Kapazität des Personals für eine Aufgabe gebunden, die für den Prozess nicht zwingend erforderlich ist, und die ansonsten für wertvollere Aufgaben wie die ärztliche Versorgung der Patient*innen verwendet werden könnte.

Die Einbindung von Empfehlungssystemen in diesen Prozess kann dazu beitragen, diesen Prozess effizienter zu gestalten. Dafür können sie in zwei Prozessschritten eingesetzt werden. Erstens als Hilfsmittel für das medizinische Fachpersonal, was den Prozess der Verfeinerung von Anfragen effizienter macht. Zweitens als Hilfsmittel für die anfordernden Ärzt*innen, was es ihnen ermöglicht, Diagnosen feingranularer und präziser anzufordern, ohne sie dabei mit einer Vielzahl möglicher Optionen und einem zeitaufwändigen Auswahlprozess zu belasten.

Neben diesen Prozessschritten sind auch andere technische Überlegungen für die Gestaltung des Empfehlungssystems und für die sich daraus ergebenden möglichen ethischen Überlegungen von Bedeutung. So stellt sich beispielsweise die Frage, auf welcher Datenbasis das Empfehlungssystem trainiert werden soll. Durchgeführte und abgerechnete Diagnosen werden hier seit langem erfasst, da sie aus buchhalterischen Gründen digital dokumentiert werden müssen. Sie stellen eine relativ saubere Datenbasis dar, da von der Beantragung bis zur Durchführung der Untersuchung mehrere Expert*innen beteiligt waren, mögliche Fehler zu korrigieren. Nachteilig ist, dass viele die-

ser Aufzeichnungen nicht mit den ursprünglichen Anträgen und den ihnen zugrunde liegenden Entscheidungen in Verbindung gebracht werden können, da sie möglicherweise in der Vergangenheit telefonisch gestellt wurden. Außerdem spiegeln sie nicht die von der anfragenden Stelle bevorzugte Granularität wider, da sie hauptsächlich der Abrechnung der Leistungen dienen. Auf der anderen Seite sind die auf der anfordernden Seite verfügbaren Daten meist unstrukturiert und auf recht grobe Diagnosekategorien beschränkt. Zudem ist die Qualität der Daten nicht gesichert, da beispielsweise relevante Anteile der Anfragen von unerfahrenen Ärzt*innen generiert werden können, die in nachfolgenden Prozessschritten von Expert*innen für die jeweiligen Diagnosen verfeinert werden können. Dies wird auch im Hinblick auf den Recommender-Ansatz relevant, da zum Beispiel Ansätze, die auf Collaborative Filtering basieren, unter der Datenqualität leiden könnten und dadurch möglicherweise nicht optimale Untersuchungen empfehlen. Dies erfordert wiederum eine nachträgliche Verfeinerung der Anfragen, wohingegen wissensbasierte Ansätze, die auf klinischen Leitlinien basieren, in ähnlichen Anwendungen nachweislich bessere Ergebnisse liefern (Mei et al. 2015).

Unter dem Gesichtspunkt der Akzeptanz bringt dies ebenfalls Herausforderungen mit sich, da sich die anfordernden Ärzt*innen durch Vorschläge, die über den ursprünglich beabsichtigten Rahmen hinausgehen, bevormundet fühlen könnten, während dies andererseits das Potenzial bietet, die Qualität der Anfragen zu verbessern und die Anzahl der erforderlichen Nachbesserungen zu verringern. Das Anbieten von Empfehlungen, die die Intention des Anfragenden zu genau modellieren, wirft dagegen Fragen der (gefühlten) Verantwortlichkeit auf, da die anfragenden Ärzt*innen in Verhaltensmuster verfallen könnten, die dazu führen, sich zu sehr auf die Empfehlung zu verlassen.

Neben diesen Überlegungen wirft die Entwicklung und der Einsatz eines Empfehlungssystems in diesem sensiblen medizinischen Kontext auch mehrere Fragen in Hinblick auf Fairness, Verantwortlichkeit und Transparenz auf, die in den folgenden Abschnitten behandelt werden.

3.2 Verantwortlichkeit

Die im vorangegangenen Abschnitt skizzierten Aspekte der Verantwortlichkeit für ein Empfehlungssystem können aus zwei Perspektiven mit auf den ersten Blick widersprüchlichen Sichtweisen betrachtet werden. Aus praktischer Sicht entlastet ein Empfehlungssystem, das die Grundlage für die Beantragung der gewünschten Untersuchung bietet, das Assistenzpersonal mit dem Potenzi-

al, es in der Prozesskette letztlich einzusparen. Aus sozio-ökonomischer Sicht geht es um die Verantwortung für diese Mitarbeiter*innen, da sich für das Krankenhaus Möglichkeiten ergeben könnten, Personal abzubauen, um Kosten zu sparen. Diese schwerwiegenden Folgen für die Betroffenen könnten gar den Einsatz als »verantwortungsvolles« Empfehlungssystem nicht mehr rechtfertigen.

Aus einer anderen Perspektive kann die gleiche Situation zu einem anderen Ergebnis führen. Wenn das Empfehlungssystem es ermöglicht, eine große Anzahl von Standardfällen direkt anzufordern, kommt die Zeitersparnis den Patient*innen zugute. Das Assistenzpersonal kann sich dann auf schwierige Fälle konzentrieren, Anfragen bearbeiten, die einer weiteren Klärung bedürfen, oder mit anderen Aufgaben betraut werden, die die Versorgung aus Sicht der Patient*innen insgesamt verbessern.

Diese Überlegungen deuten darauf hin, dass das tatsächliche Ergebnis hinsichtlich der Verantwortungsaspekte von den Entscheidungen der Krankenhausleitung abhängt, was sich mit der kürzlich von Gansky und McDonald (2022) geführten Diskussion deckt, die anmerken, dass der organisatorische Kontext, in dem das System eingesetzt werden soll, berücksichtigt werden muss.

Daher scheint der Einsatz eines Empfehlungssystems zum Wohle der Patient*innen, die in einer Krankenhausumgebung im Mittelpunkt stehen, vertretbar, wenn Compliance- und Sicherheitsaspekte eingehalten werden.

3.3 Fairness

Fairness von Empfehlungssystemen kann aus einer prozeduralen Perspektive (Lee et al. 2019) mit einem Fokus auf Fairness innerhalb des Entscheidungsprozesses oder aus einer ergebnisorientierten Perspektive der Behandlung ähnlicher Individuen oder Gruppen (Biega, Gummadi und Weikum 2018) bewertet werden. Während Fairness im Allgemeinen mit der Verringerung oder Vermeidung von Verzerrungen im Entscheidungsprozess des maschinellen Lernens verbunden ist, die oft durch die Daten eingeführt werden (Zou und Schiebinger 2018), kann sie aus einer prozeduralen Perspektive (Lee et al. 2019) mit einem Fokus auf Fairness innerhalb des Entscheidungsprozesses oder aus einer ergebnisorientierten Perspektive bei der Behandlung ähnlicher Individuen oder Gruppen (Biega, Gummadi und Weikum 2018) angegangen werden.

Der Unterschied zwischen einer prozeduralen und einer ergebnisorientierten Perspektive kann zu unterschiedlichen Bewertungen in Bezug auf ethi-

sche Belange und somit zu unterschiedlichen Ergebnissen führen. Tatsächlich kann verfahrensbezogene Fairness zu Ergebnissen führen, die als ungerecht empfunden werden, und umgekehrt (Lee et al. 2019). Speziell im Bereich der Gesundheitsversorgung wurde eine ähnliche Frage von Tsuchiya und Dolan (2009) untersucht, die die Sichtweise der Öffentlichkeit auf ergebnis- und gewinnbezogene Fairness betrachtet haben. Sie warfen Fragen bezüglich der Verzerrung der wahrgenommenen Fairness in der Gesellschaft auf. In ihrer Studie zeigen Tsuchiya und Dolan (2009), dass die Mehrheit der Befragten in Bezug auf den sozialen Status eine ergebnisorientierte Fairness im medizinischen Kontext bevorzugt, während eine beträchtliche Minderheit eine gewinnorientierte Fairness bevorzugt. Für die ergebnisorientierte Perspektive ist jedoch die Definition des Ergebnisses von großer Bedeutung. Offensichtlich kann die Empfehlung einer Untersuchung nicht als Ergebnis angesehen werden, da unterschiedliche (potenziell geschützte) Untergruppen, beispielsweise Kinder und ältere Menschen, aus medizinischer Sicht unterschiedliche Untersuchungen benötigen. Die Definition des Ergebnisses ist auch für die Bewertung der Empfehlung selbst von Bedeutung, da das Lernen aus in der Vergangenheit durchgeführten Untersuchungen als Grundwahrheit nicht unbedingt die beste Wahl für die Patient*in widerspiegelt und den Erfolg nachfolgender Behandlungen nicht bewerten kann. Eine Lösung, wie sie von Mei et al. (2015) vorgeschlagen wurde, ist die Verwendung einer kohortenstudienbasierten Auswertung zur Bewertung des Erfolgs der empfohlenen Behandlungen und zur Bewertung der Fairness. In der hier betrachteten Fragestellung ist das Ergebnis jedoch schwer zu definieren, da selbst einfache Kriterien, wie die Wartezeit bis zur Diagnosestellung, durch gruppenspezifische medizinische Aspekte, wie beispielsweise die Dringlichkeit, verzerrt werden können und sich diese Verzerrung auch in den Daten wiederfindet. Zusätzlich zu solchen gerechtfertigten und gewünschten Verzerrungen können sich in den Daten aber auch implizite Verzerrungen wiederfinden, die im Sinne eines fairen Entscheidungssystems nicht durch das System übernommen werden sollten.

Dies zeigt, dass bestimmte Kompromisse in Bezug auf Fairness und Ethik aus fachspezifischer Sicht notwendig sind, um die medizinische und ökonomische Realität widerzuspiegeln, wie sie in der täglichen medizinischen Praxis benötigt wird (Beauchamp und Childress 1994), und um damit die Akzeptanz für alle Beteiligten zu gewährleisten (Milano, Taddeo und Floridi 2020). Dennoch muss die Prüfung dieser Kompromisse auf ihre klinische und ethische Validität in den Bewertungsprozess des Empfehlungssystems einbezogen werden.

Aus algorithmischer Sicht spielt die Unverzerrtheit der Daten in Hinblick auf Ansätze, die Collaborative Filtering verwenden, eine größere Rolle als bei Content-basierten Verfahren, insbesondere solchen, die auf medizinischem Wissen basieren, da sich potentielle Fairnessprobleme aus der Vergangenheit, die in den Daten vorkommen, durch den inhärenten Zusammenhang mit ähnlichen Fällen wiederholen würden. Zahlreiche Forschungsergebnisse versuchen außerdem auf algorithmischer Ebene unterschiedliche Perspektiven von Fairness sicherzustellen (Wang et al. 2023), die für die Umsetzung eines medizinischen Empfehlungssystems Beachtung finden sollten.

3.4 Rechenschaftspflicht

Die Rechenschaftspflicht im Zusammenhang mit Systemen zur Empfehlung von Untersuchungen konzentriert sich auf die Frage, wer für den potenziellen Schaden, den das komplexe sozio-technische System einschließlich der anfordernden Ärzt*innen, des Empfehlungssystems und seiner Entwickler*innen, des medizinischen Assistenzpersonals und der ausführenden Ärzt*innen unter Umständen verursacht, verantwortlich ist und wer sich dafür verantwortlich fühlt. In einer aktuellen Studie des Europäischen Parlamentarischen Forschungsdienstes über KI im Gesundheitswesen (Lekadir et al. 2022) wird der Bedarf an neuen Mechanismen und Rahmenbedingungen zur Gewährleistung einer angemessenen Rechenschaftspflicht bei medizinischer KI festgestellt. Lücken in der KI-Rechenschaftspflicht finden sich in drei Aspekten: Aus rechtlicher Sicht fehlende Regelungen und Definitionen für die Rechenschaftspflicht, die Schwierigkeit, Rollen und damit verbundene Verantwortlichkeiten innerhalb des soziotechnischen Prozesses zu definieren, und das Fehlen einer rechtlichen und ethischen Governance für KI-Entwicklende und -Herstellende. Während die meisten vorgeschlagenen Maßnahmen eine regulatorische Sichtweise einnehmen, schlägt eine der Maßnahmen vor, Prozesse zu implementieren, um die Rollen von KI und Nutzende zu identifizieren, wenn KI-basierte Entscheidungen Patient*innen schaden könnten und die Verantwortlichkeiten explizit zu machen. In ähnlicher Weise schlagen Habli, Lawton und Porter (2020) vor, KI-Entwicklende und -Ingenieur*innen einzubeziehen, wenn es darum geht, die Verantwortlichkeit für Entscheidungen im Rahmen des KI-gestützten Prozesses zu bewerten. Darüber hinaus erörtern sie, dass die Rechenschaftspflicht unmittelbar mit Sicherheit verbunden ist, die während der KI-Entwicklung nicht vollständig vorhersehbar

ist, insbesondere im Hinblick auf mögliche Verhaltensanpassungen von Ärzt*innen, Patient*innen und dem System selbst.

Da die Ärzt*innen die endgültige Entscheidung über die Annahme oder Ablehnung von Empfehlungen treffen, schlagen wir vor, das Bewusstsein für Verantwortlichkeit und Sicherheitsbedenken als Teil des Anforderungsprozesses zu schärfen, indem die Annahmequote der vorgeschlagenen Empfehlung beobachtet wird. Wenn diese Auswertung darauf hindeutet, dass sich die anfragende Person zu sehr auf die Empfehlung verlässt, können Schritte zur Sensibilisierung eingeleitet werden, beispielsweise durch die Empfehlung einer unzulässigen Diagnose, die, wenn sie unreflektiert akzeptiert wird, mit einer Warnmeldung gestoppt wird.

3.5 Transparenz

Transparenz wird von Smith (2021) als Schlüsselaspekt für die Rechenschaftspflicht und damit für die Akzeptanz identifiziert, die argumentiert, dass Ärzt*innen für den Einsatz in der klinischen Praxis über ihre Entscheidung Rechenschaft ablegen müssen und intransparente KI-Systeme ablehnen werden, da sie deren Ergebnis nicht nachvollziehen können. Andererseits behaupten Clement, Ren und Curley (2021) in ihrer empirischen Studie, dass Transparenz die Akzeptanz minderwertiger Empfehlungen begünstigt, also eine Verhaltensänderung bewirkt, indem sie potentiell falsche Modellempfehlungen mit plausibel klingenden Erklärungen versieht.

Da die von den Ärzt*innen formulierten Anforderungen auf medizinischen Entscheidungen beruhen, dient das Empfehlungssystem in unserem Szenario in erster Linie dazu, den Auswahlprozess und nicht den Entscheidungsprozess zu verbessern. Daher müssen die Ärzt*innen in diesem Rahmen nicht für das Ergebnis des Systems, also die Empfehlung, die Verantwortung übernehmen, sondern für ihre Entscheidung. Eine von der eigentlichen Absicht abweichende Empfehlung muss einer medizinischen Bewertung unterzogen werden, bevor sie als praktikable Option eingestuft wird. Eine automatisierte Rationalisierung oder Erklärung der Empfehlung könnte daher den Bewertungsprozess abkürzen und damit die notwendige Sorgfalt untergraben, was auch mit der allgemeineren Beobachtung der Aufwertung algorithmischer Entscheidungen in Zusammenhang steht (Logg, Minson und Moore 2019). Eine weniger eingreifende Form der Transparenz, die Ärzt*innen die notwendigen Informationen liefert, ohne die eigentliche Argumentation vorwegzunehmen, kann daher besser geeignet sein. Die Bereitstellung

ähnlicher Fälle zum Vergleich für einen auf Collaborative Filtering basierenden Ansatz oder die Unterstützung wissensbasierter Empfehlungen durch klinische Leitlinien könnten ein vielversprechender Kompromiss zwischen der Undurchsichtigkeit der Empfehlungen und feinkörnigen Begründungen darstellen, ohne zu unerwünschten Verhaltensänderungen zu führen.

3.6 Compliance

Was die Einhaltung der Vorschriften betrifft, so gelten in Deutschland und der Europäischen Union für medizinische KI die 2017/746 In Vitro Diagnostic Medical Devices Regulation (IVDR) und die 2017/745 Medical Devices Regulations (MDR) (Lekadir et al. 2022), wobei letztere eher für die Festlegung von Diagnoseempfehlungen gilt. Kiseleva (2020) kommt zu dem Schluss, dass diese Verordnung als anfänglicher rechtlicher Rahmen dienen kann, jedoch in Bezug auf Transparenz und Rechenschaftspflicht erweitert werden muss. Dies wird durch die Notwendigkeit einer Risikobewertung im Vorschlag des Forschungsdienstes des Europäischen Parlaments über Künstliche Intelligenz im Gesundheitswesen (Lekadir et al. 2022) ergänzt, der in Übereinstimmung mit dem Vorschlag der Europäischen Kommission für eine KI-Regulierung zusätzlich vorschlägt, Anwendungen nach Risikostufen von inakzeptabel bis minimal einzustufen. Für Empfehlungssysteme wurde die Anwendbarkeit dieser Regelung kürzlich von Schwemer (2021) diskutiert, wobei die Anwendung im medizinischen Kontext als potenzielle Hochrisikoanwendungen eingestuft werden dürften (Lekadir et al. 2022). Aus praktischer Sicht bietet die Initiative FUTURE-AI (Lekadir et al. 2021) Leitlinien und Best Practices für vertrauenswürdige KI in der Medizin, die bei der Gestaltung eines Empfehlungssystems berücksichtigt werden sollten.

3.7 Sicherheit

Der Aspekt der Sicherheit ist nicht nur mit der eigentlichen Performance des Empfehlungssystems verbunden, was die offensichtliche Fragestellung darstellt, wie geeignet die Empfehlungen sind und ob der Einsatz des Systems positive Auswirkungen hat, oder ob falsche oder ungenaue Empfehlungen eines schwachen Empfehlungssystems einen negativen Einfluss haben können, etwa durch zusätzlichen Zeitaufwand bei der Verfeinerung der Anfrage bis hin zu nicht optimalen diagnostischen Entscheidungen, die wichtige Gesundheitszustände übersehen und somit direkt oder indirekt Schaden

für die Patient*innen verursachen (Lekadir et al. 2022). Stattdessen kann die Sicherheit auch aus einer nutzungszentrierten Perspektive angegangen werden, wie im Zusammenhang der Rechenschaftspflicht erörtert, wo der Missbrauch eines Systems zu Sicherheitsproblemen führen kann, sowie aus einer datenorientierten Perspektive, beispielsweise, wenn die Daten falsche Informationen oder Rauschen enthalten oder die Datenpunkte aus einer auf klinischen Leitlinien basierenden Perspektive veraltet sind. Dies spiegelt sich auch in praktischen Richtlinien und Vorschriften wider, da Sicherheit am besten ganzheitlich betrachtet wird, von der Datenerfassung, der Annotation, über das Systemdesign und der Evaluierung bis hin zum sozio-technischen und organisatorischen Kontext, in dem es verwendet wird (Lekadir et al. 2021; Dobbe 2022) und in diesem Komplex entsprechend geprüft wird (Falco et al. 2021).

Für die Bewertung der Sicherheit unseres Empfehlungssystems bedeutet dies, dass wir uns nicht nur auf Leistungskennzahlen verlassen können, zumal die Daten selbst anfällig für Fehler und Rauschen sein können. Während die Leistung des Systems in einem Bereich liegen muss, in dem der potenzielle Nutzen die Sicherheitsrisiken überwiegt, um brauchbar zu sein, muss der Nutzen und die Sicherheit aus einer Ergebnisperspektive konsequent überwacht werden, insbesondere während des Betriebs.

3.8 Zusammenfassung und Diskussion

Unter Berücksichtigung der vorangegangenen Überlegungen, die wir im Folgenden nochmals zusammenfassen, könnte der Einsatz eines verantwortungsvollen Empfehlungssystems zum Wohle der Patient*innen vertretbar sein, wenn Compliance- und Sicherheitsaspekte eingehalten werden können. Hinsichtlich Fairness scheint das Ergebnis in der hier betrachteten Fragestellung schwer bewertbar zu sein, da selbst einfache Kriterien wie die Wartezeit bis zur Diagnosestellung durch gruppenspezifische medizinische Aspekte, wie beispielsweise die Dringlichkeit, verzerrt werden können. Dies zeigt, dass bestimmte Kompromisse in Bezug auf Fairness und Ethik aus fachspezifischer Sicht notwendig sind, um die medizinische und ökonomische Realität widerzuspiegeln (Beauchamp und Childress 1994) und damit die Akzeptanz für alle Beteiligten zu gewährleisten (Milano, Taddeo und Floridi 2020). Daher muss die Prüfung dieser Kompromisse auf ihre klinische und ethische Validität in den Bewertungsprozess des Empfehlungssystems einbezogen werden. Die Algorithmik des Empfehlungssystems ist nicht in der Lage, diese

Balance automatisch zu gewährleisten und bedarf daher Anpassungen, falls die Fairness nicht sichergestellt ist.

Da die Ärzt*innen die endgültige Entscheidung über die Annahme oder Ablehnung von Empfehlungen treffen, sind sie auch verantwortlich und damit rechenschaftspflichtig. Wir schlagen deshalb vor, das Bewusstsein für Verantwortlichkeit und Sicherheitsbedenken als Teil des Anforderungsprozesses immer wieder zu schärfen, indem beispielsweise die Annahmequote der vorgeschlagenen Empfehlung beobachtet und hinterfragt wird. Hinsichtlich der Transparenz könnten plausibel klingende Erklärungen die nötige Sorgfalt, diese nicht unhinterfragt zu übernehmen, einschränken. Eine weniger eingreifende Form der Transparenz, die Ärzt*innen die notwendigen Informationen liefert, ohne die eigentliche Argumentation vorwegzunehmen, dürfte daher besser geeignet sein, um trotzdem Transparenzanforderungen an das Systems zu adressieren. Nach dem Vorschlag des Forschungsdienstes des Europäischen Parlaments über Künstliche Intelligenz im Gesundheitswesen (Lekadir et al. 2022) dürfte die Anwendung von Empfehlungssystemen im medizinischen Kontext als potenzielle Hochrisikoanwendungen eingestuft werden (Lekadir et al. 2022). Aus praktischer Sicht bietet die Initiative FUTURE-AI (Lekadir et al. 2021) Leitlinien und Best Practices für vertrauenswürdige KI in der Medizin, die bei der Gestaltung eines Empfehlungssystems berücksichtigt werden sollten, um den hohen Anforderungen als Hochrisikoanwendung gerecht zu werden. Sicherheitsaspekte sollten am besten ganzheitlich betrachtet werden, von der Datenerfassung, der Annotation, über das Systemdesign und der Evaluierung bis hin zum sozio-technischen und organisatorischen Kontext, in dem es verwendet wird (Lekadir et al. 2021; Dobbe 2022) und in diesem Komplex entsprechend geprüft werden (Falco et al. 2021), insbesondere auch während des Betriebs.

4. Fazit

In dieser Arbeit haben wir ein Empfehlungssystem für medizinische Untersuchungen in einem anwendungsorientierten Kontext skizziert. Wir haben die Implikationen der Designentscheidungen, des Anwendungsfalls und des Systems im Hinblick auf Verantwortung, Fairness, Rechenschaftspflicht, Transparenz, Compliance und Sicherheit erörtert und einen Überblick über die Möglichkeiten gegeben, wie diese Implikationen und Probleme aus praktischer Sicht adressiert werden können. Im Kontext unseres Empfeh-

lungssystems konnten wir damit wichtige ethische Aspekte, die für eine verantwortungsvolle Gestaltung und Implementierung relevant sind, operationalisieren und im Kontext bisheriger Arbeiten bewerten, und hoffen, damit Impulse für die weitere Entwicklung eines solchen Systems und deren Umsetzung in der Praxis zu geben.

Danksagung

Diese Forschung wurde durch das Bayerische Staatsministerium für Wissenschaft und Kunst im Projekt »Digitalisierungszentrum für Präzisions und Telemedizin« (DZ.PTM) im Rahmen des Masterplans »BAYERN DIGITAL II« gefördert.

Literatur

Adomavicius, Gediminas, and Alexander Tuzhilin. 2005. »Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions«. *IEEE transactions on knowledge and data engineering* 17 (6): 734–749.

Aggarwal, Charu C., et al. 2016. *Recommender systems*. Bd. 1. Springer.

Balabanović, Marko, and Yoav Shoham. 1997. »Fab: content-based, collaborative recommendation«. *Communications of the ACM* 40 (3): 66–72.

Beauchamp, Tom L. and James F. Childress 1994. Principles of biomedical ethics. Oxford University Press.

Biega, Asia J., Krishna P. Gummadi and Gerhard Weikum. 2018. »Equity of attention: Amortizing individual fairness in rankings«. In *The 41st international acm sigir conference on research & development in information retrieval*, 405–414.

Bobadilla, Jesús, Fernando Ortega, Antonio Hernando and Abraham Gutiérrez. 2013. »Recommender systems survey«. *Knowledge-based systems* 46: 109–132.

Burke, Robin. 2002. »Hybrid recommender systems: Survey and experiments«. In *User Modeling and User-Adapted Interaction* (12): 331–370.

Chen, Jiawei, Hande Dong, Xiang Wang, Fuli Feng, Meng Wang and Xiangnan He. 2023. »Bias and debias in recommender system: A survey and future directions«. *ACM Transactions on Information Systems* 41 (3): 1–39.

Clement, Jeffrey, Yuqing Ching Ren and Shawn Curley. 2021. »Increasing System Transparency About Medical AI Recommendations May Not Improve Clinical Experts' Decision Quality«. Available at SSRN 3961156.

Dallmann, Alexander, Johannes Kohlmann, Daniel Zoller and Andreas Hotho. 2021. »Sequential Item Recommendation in the MOBA Game Dota 2«. In *2021 International Conference on Data Mining Workshops (ICDMW)*, 10–17. IEEE.

Dobbe, Roel. 2022. »System Safety and Artificial Intelligence«. In *2022 ACM Conference on Fairness, Accountability, and Transparency*, 1584–1584.

Ekstrand, Michael D., Robin Burke and Fernando Diaz. 2019. »Fairness and discrimination in recommendation and retrieval«. In *Proceedings of the 13th ACM Conference on Recommender Systems*, 576–577.

FAccTRec. 2022. *RecSys – ACM Recommender Systems*. Accessed February 20, 2023. https://recsys.acm.org/recsys22/facctrec/.

Falco, Gregory, Ben Shneiderman, Julia Badger, Ryan Carrier, Anton Dahbura, David Danks, Martin Eling, Alwyn Goodloe, Jerry Gupta, Christopher Hart et al. 2021. »Governing AI safety through independent audits«. *Nature Machine Intelligence* 3 (7): 566–571.

Fischer, Elisabeth, Daniel Zoller and Andreas Hotho. 2021. »Comparison of Transformer-Based Sequential ProductRecommendation Models for the Coveo Data Challenge«. *SIGIR Workshop On eCommerce* (July).

Fox, Sarah, Christina Harrington, Aziz Huq and Chenhao Tan, (Hg.). 2023. *FAccT '23: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. Chicago, IL, USA: Association for Computing Machinery.

Gansky, Ben, and Sean McDonald. 2022. »CounterFAccTual: How FAccT Undermines Its Organizing Principles«. *In 2022 ACM Conference on Fairness, Accountability, and Transparency*, 1982–1992. FAccT '22. Seoul, Republic of Korea: Association for Computing Machinery. DOI: https://doi.org/10.1145/3531146.3533241.

Gomez-Uribe, Carlos A., and Neil Hunt. 2015. »The netflix recommender system: Algorithms, business value, and innovation«. *ACM Transactions on Management Information Systems (TMIS)* 6 (4): 1–19.

Habli, Ibrahim, Tom Lawton and Zoe Porter. 2020. »Artificial intelligence in health care: accountability and safety«. *Bulletin of the World Health Organization* 98 (4): 251.

Jäschke, Robert, Leandro Balby Marinho, Andreas Hotho, Lars Schmidt-Thieme and Gerd Stumme. 2007. »Tag Recommendations in Folksonomies«. In *Knowledge Discovery in Databases: PKDD 2007. 11th European*

Conference on Principles and Practice of Knowledge Discovery in Databases, Warsaw, Poland, September 17–21, 2007, Proceedings, edited by Joost N. Kok, Jacek Koronacki, Ramon López de Mántaras, Stan Matwin, Dunja Mladenic and Andrzej Skowron, 4702: 506–514. Lecture Notes in Computer Science. Springer. https://www.kde.cs.uni-kassel.de/wp-content/uploads/hotho/pub/2007/kdml_recommender_final.pdf.

Kiseleva, Anastasiya. 2020. »AI as a Medical Device: Is It Enough to Ensure Performance Transparency and Accountability in Healthcare?« *European Pharmaceutical Law Review*, Nr. 1.

Lee, Min Kyung, Anuraag Jain, Hea Jin Cha, Shashank Ojha and Daniel Kusbit. 2019. »Procedural justice in algorithmic fairness: Leveraging transparency and outcome control for fair algorithmic mediation«. *Proceedings of the ACM on Human-Computer Interaction* 3 (CSCW): 1–26.

Lekadir, Karim, Richard Osuala, Catherine Gallin, Noussair Lazrak, Kaisar Kushibar, Gianna Tsakou, Susanna Aussó, Leonor Cerdá Alberich, Kostas Marias, Manolis Tsiknakis et al. 2021. »FUTURE-AI: Guiding Principles and Consensus Recommendations for Trustworthy Artificial Intelligence in Medical Imaging«. *arXiv preprint arXiv*: 2109.09658.

Lekadir, Karim, Gianluca Quaglio, Anna Tselioudis Garmendia and Catherine Gallin. 2022. *Principles of biomedical ethics*. European Parliamentary Research Service.

Logg, Jennifer M, Julia A. Minson and Don A Moore. 2019. »Algorithm appreciation: People prefer algorithmic to human judgment«. *Organizational Behavior and Human Decision Processes* 151: 90–103.

Mehrabi, Ninareh, Fred Morstatter, Nripsuta Saxena, Kristina Lerman and Aram Galstyan. 2021. »A survey on bias and fairness in machine learning«. *ACM Computing Surveys (CSUR)* 54 (6): 1–35.

Mei, Jing, Haifeng Liu, Xiang Li, Yiqin Yu and Guotong Xie. 2015. »Outcomedriven Evaluation Metrics for Treatment Recommendation Systems.« In *MIE*, 190–194.

Milano, Silvia, Mariarosaria Taddeo and Luciano Floridi. 2020. »Recommender systems and their ethical challenges«. *Ai & Society* 35 (4): 957–967.

Mooney, Raymond J., and Loriene Roy. 2000. »Content-based book recommending using learning for text categorization«. In *Proceedings of the fifth ACM conference on Digital libraries*, 195–204.

Nguyen, Tien T., Pik-Mai Hui, F. Maxwell Harper, Loren Terveen and Joseph A. Konstan. 2014. »Exploring the filter bubble: the effect of using recom-

mender systems on content diversity«. In *Proceedings of the 23rd international conference on World wide web*, 677–686.

Pazzani, Michael J. 1999. »A framework for collaborative, content-based and demographic filtering«. *Artificial intelligence review* 13: 393–408.

Pazzani, Michael J, and Daniel Billsus. 2007. »Content-based recommendation systems«. *The adaptive web: methods and strategies of web personalization*, 325–341.

Schein, Andrew I., Alexandrin Popescul, Lyle H. Ungar and David M. Pennock. 2002. »Methods and metrics for cold-start recommendations«. In *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*, 253–260.

Schwemer, Sebastian Felix. 2021. »Recommender Systems in the EU: from Responsibility to Regulation?« In *FAccTRec Workshop*, Bd. 21.

Shi, Yue, Martha Larson and Alan Hanjalic. 2014. »Collaborative filtering beyond the user-item matrix: A survey of the state of the art and future challenges«. *ACM Computing Surveys (CSUR)* 47 (1): 1–45.

Smith, Helen. 2021. »Clinical AI: opacity, accountability, responsibility and liability«. *AI & SOCIETY* 36 (2): 535–545.

Song, Yading, Simon Dixon and Marcus Pearce. 2012. »A survey of music recommendation systems and future perspectives«. In *9th international symposium on computer music modeling and retrieval*, 4: 395–410.

Su, Xiaoyuan, and Taghi M. Khoshgoftaar. 2009. »A survey of collaborative filtering techniques«. *Advances in artificial intelligence* 2009.

Sun, Leilei, Chuanren Liu, Chonghui Guo, Hui Xiong and Yanming Xie. 2016. »Data-driven automatic treatment regimen development and recommendation«. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 1865–1874.

Tsuchiya, Aki, and Paul Dolan. 2009. »Equality of what in health? Distinguishing between outcome egalitarianism and gain egalitarianism«. *Health economics* 18 (2): 147–159.

Wang, Yifan, Weizhi Ma, Min Zhang, Yiqun Liu and Shaoping Ma. 2023. »A Survey on the Fairness of Recommender Systems«. *ACM Transactions on Information Systems* 41, Nr. 3 (July): 1–43. DOI: https://doi.org/10.1145/3547333.

Zhang, Shuai, Lina Yao, Aixin Sun and Yi Tay. 2019. »Deep learning based recommender system: A survey and new perspectives«. *ACM computing surveys (CSUR)* 52 (1): 1–38.

Zou, James, and Londa Schiebinger. 2018. »AI can be sexist and racist—it's time to make it fair.« *Nature* 559 (7714): 324–326.