21. Wahlperiode

Deutscher Bundestag
Innenausschuss

Ausschussdrucksache 21(4)258 E
vom 20. November 2026

Schriftliche Stellungnahme

von Prof. Dr. Alexander Steen, Gesellschaft für Informatik, Universität Greifswald vom 19. November 2026

zu dem

Gesetzentwurf der Bundesregierung
Entwurf eines Gesetzes zum Einsatz künstlicher Intelligenz in der Migrationsverwaltung
BT-Drucksache 21/8275

Dem Ausschuss ist das vorliegende Dokument in nicht barrierefreier Form zugeleitet worden.

FACHBEREICH
KÜNSTLICHE
INTELLIGENZ

UNIVERSITÄT GREIFSWALD
Wissen lockt. Seit 1456

Universität Greifswald, Institut für Mathematik und Informatik, Walther-Rathenau-Straße 47, 17489 Greifswald

Innenausschuss des Deutschen Bundestages
11011 Berlin

Mathematisch-Naturwissenschaftliche Fakultät

Institut für Mathematik u. Informatik
Prof. Dr. Alexander Steen
Walther-Rathenau-Straße 47
17489 Greifswald
Telefon: +49 3834 420 4634
alexander.steen@uni-greifswald.de

Co-Sprecher des Fachbereichs KI der Gesellschaft für Informatik

20. November 2026

Stellungnahme zum *Entwurf eines Gesetzes zum Einsatz künstlicher Intelligenz in der Migrationsverwaltung, BT-Drucksache 21/8275*

Kurzzusammenfassung

A. Eine sinnvolle Digitalisierung und Optimierung von Verfahrensdauern ist begrüßenswert.

B. Der Nutzen automatisierter Anwendungen muss anhand ihrer tatsächlichen Wirkung im sozio-technischen Kontext konkreter Verwaltungsverfahren untersucht und beurteilt werden.

C. Die Entwicklungsbefugnisse bleiben hinsichtlich der vorgesehenen Aufgaben und erforderlichen Daten zu unbestimmt und sollten durch konkrete zulässige und ausgeschlossene Verwendungszwecke begrenzt werden.

D. Allgemeine Qualitätszusicherungen und Diskriminierungsverbote erfordern überprüfbare, anwendungsbezogene Anforderungen und dokumentierte Nachweise über deren Erfüllung.

E. Die Auffindbarkeit von Internetinhalten belegt weder ihre Authentizität und sachliche Richtigkeit noch ihre Zuordnung zur betroffenen Person.

F. Einer nachteiligen Verwendung von Informationen aus dem Internet muss eine belastbare Prüfung ihrer Herkunft, Personenzuordnung und konkreten Aussagekraft vorausgehen.

G. Menschliche Entscheidungsverantwortung und verständliche KI-Erklärungen können alleine die unreflektierte Übernahme fehlerhafter automatisierter Empfehlungen nicht verhindern.

H. Internetrecherchen können unbeabsichtigt sensible Zusammenhänge gegenüber Drittanbietern offenlegen und erlauben diesen prinzipiell eine gezielte Beeinflussung der angezeigten Suchergebnisse.

I. Vermeidbare Personenbezüge schaffen beim Verfahrensmonitoring erhebliche Datenschutzrisiken.

J. Ohne gemeinsame Referenzmaßstäbe für automatisierte Anwendungen drohen Insellösungen, die dem Ziel einer einheitlicheren Verwaltungspraxis entgegenwirken.

K. Die Beschaffung von Software und die Nutzung externer Dienstleistungen können neue Abhängigkeiten schaffen und die digitale Souveränität der Behörden gefährden.

L. Die erwarteten Verbesserungen können nur eintreten, wenn Entwicklung, Einführung und Betrieb ausreichend finanziert werden.

1. Vorbemerkungen

Die nachfolgende Stellungnahme bezieht sich auf den Entwurf eines Gesetzes zum Einsatz künstlicher Intelligenz in der Migrationsverwaltung der Bundesregierung (BT-Drs. 21/8275). Sie gibt die persönliche technische Facheinschätzung des Autors wider, nicht notwendigerweise die des Fachbereichs Künstliche Intelligenz (FBKI) der Gesellschaft für Informatik (GI).

Das Ziel des Entwurfs, Verfahren in der Migrationsverwaltung zu beschleunigen und ihre Qualität und Einheitlichkeit zu erhöhen, ist begrüßenswert und einsichtig. Digitale Verfahren und Unterstützungssysteme können hierzu einen nennenswerten Beitrag leisten. So lässt sich beispielsweise durch die Analyse dokumentierter Bearbeitungsschritte untersuchen, an welchen Stellen Anträge regelmäßig für längere Zeit nicht weiter bearbeitet werden oder welche Prüfungen und Einzelschritte besonders viele Ressourcen beanspruchen. In so einem Fall kann anschließend untersucht werden, ob eine vorgenommene Verfahrensanpassung tatsächlich die Bearbeitungszeit verkürzt. Bei anderen möglichen Anwendungen automatisierter Verfahren kann der Nutzen wesentlich schwieriger zu bestimmen sein: Ein automatisierter Internetabgleich kann die Suche nach Informationen beschleunigen, aber ob dies auch die Sachverhaltsaufklärung verbessert, hängt unter anderem davon ab, wie zuverlässig ein Rechercheergebnis der richtigen Person zugeordnet werden kann und wie aufwändig die Prüfung der Richtigkeit und der Relevanz gefundener Informationen in der Praxis ist (BT-Drs. 21/8275, S. 1, 6-7 und 20-21).

Aus informatischer Sicht sind solche automatisierten Anwendungen insbesondere im Entwurfskontext notwendigerweise als sozio-technische Systeme aufzufassen. Mögliche positive wie auch negative emergente Eigenschaften, die aus der Interaktion des technischen Systems mit Menschen entstehen, sind hierbei ausdrücklich zu berücksichtigen. Ihre Wirkung ergibt sich aus dem Zusammenspiel von Softwareanwendungen, Daten, sachbearbeitenden Personen und den gegebenen Rahmenbedingungen und Prozessen. Bei einer entsprechenden Bewertung des Einsatzzweckes sind also stets der institutionelle Kontext sowie die von den Verfahren betroffenen Menschen zu betrachten[1]. Ob der Einsatz automatisierter Verfahren die Prozesse verbessert, hängt wesentlich von der Art der Aufgabe ab[2].

Für die Beurteilung des Entwurfs ist daher maßgeblich, ob die vorgesehene Unterstützung die Arbeitsabläufe der Verwaltungsbehörden in konkreten Fachverfahren nachweisbar verbessert. Dazu gehören die sachliche Qualität der Bewertung, die Bearbeitungszeit inklusive notwendiger manueller Nachprüfungen sowie die Möglichkeit, Fehler rechtzeitig zu erkennen und zu berichtigen. Daraus ergeben sich Anforderungen sowohl an den Schutz vor unbefugtem Zugriff und gezielter Manipulation (*Security*) als auch an die Vermeidung von Schäden, die selbst bei bestimmungsgemäßer Nutzung durch z.B. fehlerhafte Ausgaben entstehen können (*Safety*). Der Gesetzgeber kann hier die Wahl der technischen Mittel offenlassen, zugleich aber behördenübergreifend festlegen, welche Aufgaben automatisiert unterstützt werden dürfen, welche Aufgaben explizit nicht unterstützt werden dürfen und welche Qualitätsnachweise vor dem Einsatz der Systeme erforderlich sind.

Die nachfolgenden Ausführungen adressieren primär die vorgeschlagenen Änderungen des Asylgesetzes in Artikel 1 des Entwurfs. Soweit Artikel 2 für das Aufenthaltsgesetz inhaltlich entsprechende Regelungen vorsieht, sind die Ausführungen analog anwendbar.

2. Anwendungsbereich digitaler Verfahren, einschließlich KI-Systemen

Der Gesetzentwurf unterscheidet zwischen der Verarbeitung personenbezogener Daten für die Entwicklung automatisierter Anwendungen und dem Einsatz solcher Anwendungen für bestimmte Aufgaben. Automatisierte Anwendungen sind dabei nicht mit KI-Systemen gleichzusetzen: Viele Aufgaben können bereits durch klassische Software unterstützt werden.

§ 7 Abs. 2a AsylG-E erlaubt den beteiligten Behörden, die von ihnen erhobenen Daten zur Entwicklung von automatisierten Anwendungen (einschließlich KI-Systemen) zu verarbeiten, soweit dies zur Aufgabenerfüllung der jeweiligen Stelle dient. Dies schließt insbesondere das Training, die Validierung und das Testen ein (BT-Drs. 21/8275, S. 5-6 und 14-16). Ob und in welchem Umfang die Behörde das entwickelte System tatsächlich einsetzen darf, ergibt sich aus der Entwicklungsbefugnis für sich genommen nicht.

1 Siehe *Selbst et al., Fairness and Abstraction in Sociotechnical Systems, ACM, 2019. DOI: 10.1145/3287560.3287598*

2 Siehe *Vaccaro et al., When combinations of humans and AI are useful: A systematic review and meta-analysis. Nat Hum Behav 8, 2024. DOI: 10.1038/s41562-024-02024-1*

Universität Greifswald
Institut für Mathematik und Informatik
Walther-Rathenau-Straße 47, 17489 Greifswald

Prof. Dr. Alexander Steen
alexander.steen@uni-greifswald.de
Telefon: +49 3834 420 **4634**

Sekretariat: Gaby Aurell
Telefon: +49 3834 420 **4614**
gaby.aurell@uni-greifswald.de

Zwei Anwendungsfälle werden im Entwurf ausdrücklich geregelt: § 7a AsylG-E erlaubt den zuständigen Behörden, im Rahmen des Verfahrensmonitorings abgeschlossene Asylverfahren und die bisherige Entscheidungspraxis automatisiert auszuwerten. § 7b AsylG-E ermächtigt das Bundesamt zum automatisierten Abgleich mit öffentlich zugänglichen Internetdaten, wenn begründete Zweifel an relevanten Angaben bestehen (siehe Abschnitt 4 für eine nähere Auseinandersetzung zu diesem Anwendungsfall).

Konkretisierungsbedarf besteht insbesondere bei der Entwicklungsbefugnis nach § 7 Abs. 2a Satz 2 AsylG-E. Laut Entwurf müssen "die Anwendungen der Aufgabenerfüllung der zuständigen Stelle dienen und die Verarbeitung der Daten für die Entwicklung erforderlich [sein]" (BT-Drs. 21/8275, S. 5). Welche Daten tatsächlich erforderlich sind, hängt stark davon ab, welche spezifische Aufgabe das System adressieren soll. Dies lässt sich an einem Beispiel verdeutlichen, welches Bezug auf die Begründung des Entwurfs nimmt: Für die Entwicklung eines Systems, das unterschiedliche Schreibweisen eines Namens in zu prüfenden Unterlagen auffinden soll, ist die Notwendigkeit von Namensdaten nachvollziehbar. Eine Anwendung, die aus Namen auf die Herkunft oder Religion einer einzelnen Person schließen soll, hat einen wesentlich anderen Zweck. Selbst ein statistisch belegter Zusammenhang zwischen beispielsweise Namen und Herkunft in einer Vergleichsgruppe belegt allein nicht die Herkunft einer einzelnen Person, auch kann dadurch für sich genommen kein Widerspruch zu ihren Angaben begründet werden. Die vage Bezeichnung "Plausibilitätsprüfung" lässt hier einen großen, sehr relevanten Spielraum zu, welche Schlüsse eine Anwendung aus statistischen Zusammenhängen folgern soll und für welchen Zweck diese genutzt werden dürfen (BT-Drs. 21/8275, S. 16).

Dieser Spielraum ist jedoch bei der Würdigung individueller Sachverhalte und bei der Unterstützung von Ermessensentscheidungen stark zu beschränken. Soll ein System aus früheren Fallausgängen eine Bewertung für einen neuen Antrag ausgeben, bleibt offen, woran die Richtigkeit dieser Bewertung gemessen werden kann. Aus vorherigen Fallausgängen und Bescheiden allein kann nicht abgeleitet werden, ob alle relevanten Umstände eines neuen Antrags adäquat gewürdigt worden sind. Es muss vielmehr zuvor bestimmt werden, wie die Korrektheit einer Bewertung gemessen werden kann. Selbst eine hohe Übereinstimmung eines automatisierten Verfahrens mit früheren Fallentscheidungen ist daher kein hinreichender Nachweis dafür, dass das System einen neuen Einzelfall sachgerecht würdigen kann.[3]

Die Konkretisierung der vorgesehenen und zulässigen Einsatzzwecke von automatisierten Anwendungen ist auch für den wirtschaftlichen Einsatz öffentlicher Mittel relevant. Die Aufbereitung von Daten sowie die Entwicklung und das Testen einer Anwendung binden signifikante Ressourcen. Ob eine Anwendung für ihren Einsatzzweck tatsächlich geeignet ist, ist Gegenstand der Anwendungstests bzw. Erprobung. Vermeidbar ist hingegen der Aufwand für ein Vorhaben, bei dem die geplante Funktion im jeweiligen Fachverfahren generell nicht eingesetzt werden darf. Je genauer der Entwurf also die zulässigen Aufgaben beschreibt, desto verlässlicher können die beteiligten Behörden bereits bei der Auswahl von Entwicklungsvorhaben beurteilen, ob ein späterer Einsatz in Betracht kommt (vgl. BT-Drs. 21/8275, S. 13).

Empfehlung: Die zulässigen Entwicklungs- und Einsatzzwecke sollten anhand konkreter Aufgaben mit Bezug auf Fachverfahren und unter Benennung der jeweils befugten Stellen näher bestimmt werden. Explizit ausgeschlossen werden sollte insbesondere die Verwendung statistischer Gruppenmerkmale oder vergangener Fallausgänge zur automatisierten Bewertung der Glaubhaftigkeit individueller Angaben. Für zulässige Entwicklungsprojekte sollte die zuständige Behörde zudem vor der Verarbeitung personenbezogener Daten dokumentieren, welchen Verfahrensschritt die geplante Anwendung unterstützen soll, welche Ausgaben dafür erzeugt werden sollen und weshalb die vorgesehenen Arten von Daten hierfür erforderlich sind. Eine solche Regelung kann technikneutral formuliert sein und zugleich eine überprüfbare Grenze für die Datennutzung setzen.

3. Qualität und Korrektheitszusicherungen von automatisierten Anwendungen

Der Gesetzentwurf schreibt vor, dass "diskriminierende Algorithmen weder herausgebildet noch verwendet werden" dürfen (§ 7 Abs. 2a Satz 3, § 7a Abs. 3 und § 7b Abs. 2 AsylG-E). Die Begründung verweist hierzu auf technische und organisatorische Vorkehrungen, die Sicherstellung der Qualität der verwendeten Daten sowie ergänzende Anforderungen der KI-VO (BT-Drs. 21/8275, S. 16 und 19). Diese Zielsetzung ist zu unterstützen, jedoch unzureichend ausgestaltet. Für praktische Umsetzungen muss zudem spezifiziert werden, anhand welcher überprüfbaren Eigenschaften die Qualität einer automatisierten Anwendung untersucht werden soll sowie welche Nachweise und Rahmenbedingungen hierfür vorliegen müssen.

3 Siehe *Jacobs und Wallach, Measurement and Fairness, ACM, 2021. DOI: 10.1145/3442188.3445901*

Universität Greifswald
Institut für Mathematik und Informatik
Walther-Rathenau-Straße 47, 17489 Greifswald

Prof. Dr. Alexander Steen
alexander.steen@uni-greifswald.de
Telefon: +49 3834 420 **4634**

Sekretariat: Gaby Aurell
Telefon: +49 3834 420 **4614**
gaby.aurell@uni-greifswald.de

Aus softwaretechnischer Perspektive ist für eine Korrektheitszusicherung eine präzise Beschreibung des erwarteten Systemverhaltens notwendig. Auch eine korrekt implementierte Anwendung kann bei ungeeigneten Eingabedaten oder inkorrekten bzw. unvollständigen fachlichen Annahmen ungeeignete Ergebnisse liefern. Eine detaillierte Beschreibung funktionaler sowie nicht-funktionaler Anforderungen von geplanten Anwendungen ist daher vor der eigentlichen Entwicklung anzufertigen und zu validieren und ggf. im Verlauf der Entwicklung weiterzuentwickeln. Aus den Anforderungen ergeben sich dann konkrete technische Spezifikationen für die zu entwickelnde Anwendung.

Die technische Aussagekraft eines Qualitätsnachweises ist in hohem Maße von der verwendeten Prüfmethode abhängig. Softwaretests untersuchen eine Anwendung anhand ausgewählter Fälle auf Defekte im Vergleich zu dem spezifizierten erwarteten Verhalten. Durch Testen allein kann jedoch im Allgemeinen keine Fehlerfreiheit für sämtliche künftigen Fälle etabliert werden. Mit formalen Verifikationsverfahren können dagegen mathematische Nachweise für präzise beschriebene Eigenschaften unter ausdrücklich festgelegten Annahmen geführt werden. Für kritische Komponenten kann geprüft werden, ob solche formalen Nachweise mit vertretbarem Aufwand erbracht werden können. Für nachprüfbare Zusicherungen müssen daher sowohl die nachgewiesene Eigenschaft als auch die getroffenen Annahmen sowie die technischen Grenzen der Aussagekraft des Nachweises dokumentiert werden.

Bei statistischen Verfahren und aus Daten lernenden Systemen ist zu prüfen, wie aussagekräftig die gemessene Leistung für den geplanten Einsatz ist. Eine bloße hohe durchschnittliche Klassifikationsgenauigkeit kann signifikante Unterschiede zwischen verschiedenen Fallgruppen verschleiern. Es ist außerdem möglich, dass zwei Systeme auf gleichen Testdaten ähnlich gute Ergebnisse erzielen, sich aber in anderen Einsatzkontexten deutlich unterschiedlich verhalten.[4] Für die Migrationsverwaltung ergibt sich daraus die zentrale Anforderung, Qualitätsnachweise auf die konkreten Aufgaben und Fälle zu beziehen. Bei einer automatisierten Anwendung zur Zuordnung von Dokumenten wären zum Beispiel unterschiedliche Schreibweisen von Angaben, fehlende Angaben und mehrdeutige Zuordnungen gezielt zu untersuchen. Dabei sind u.a. die jeweiligen Fehlerarten, die Zahl der geprüften Fälle und die statistische Unsicherheit der Ergebnisse auszuweisen.

KI-Systeme sind dabei jedoch nicht mit Verfahren des maschinellen Lernens gleichzusetzen, auch arbeiten diese nicht notwendigerweise mit aus Daten erlernten statistischen Repräsentationen. Symbolische KI-Methoden können beispielsweise explizit formulierte Regeln und logische Zusammenhänge verarbeiten und daraus präzise Schlussfolgerungen ableiten. Für klar abgegrenzte und formal beschreibbare Kernaufgaben können solche Verfahren eingesetzt werden, deren Korrektheit bezüglich einer gegebenen Spezifikation mathematisch nachgewiesen werden kann. Bei der Auswahl technischer Verfahren sollten deshalb ausdrücklich auch symbolische Ansätze berücksichtigt werden.

Die Anforderung, diskriminierende Ergebnisse auszuschließen, bedarf einer fachlichen Konkretisierung. Es muss zum Beispiel geklärt werden, welche Fehler welche Personengruppen besonders benachteiligen könnten und anhand welcher Kriterien dies überprüft und verhindert wird. Untersuchungen machen deutlich, dass einige relevante Fairnesskriterien bei der Angabe statistischer Risikowerte im Allgemeinen nicht gleichzeitig erfüllt werden können.[5] Die Auswahl der Prüfmaßstäbe macht also eine am konkreten Verwendungszweck und den relevanten Schutzanforderungen orientierte Begründung erforderlich.

Zudem muss die Aussagekraft der Nachweise während des Betriebs erhalten bleiben. Änderungen an der Anwendung, den Fachverfahren, den ggf. verwendeten Modellen oder Datenquellen können die Voraussetzungen einer früheren Prüfung verändern. Vor geplanten technischen oder strukturellen Änderungen des digitalen Verfahrens muss deshalb sichergestellt werden, dass bisherige Nachweise weiterhin Aussagekraft besitzen.

Empfehlung: Der Entwurf sollte die allgemeinen Qualitätsanforderungen durch eine Verpflichtung zur Erstellung von anwendungsbezogenen und dokumentierten Nachweisen konkretisieren. Die zuständigen Stellen sollten bereits bei der Planung des Systems die zu prüfenden Eigenschaften, den geplanten Einsatzbereich, die Prüfmethoden und die notwendigen Kriterien für eine Freigabe festlegen. Umfang und Tiefe einer Prüfung sollten

4 Siehe *D'Amour et al., Underspecification Presents Challenges for Credibility in Modern Machine Learning, J. Mach. Learn. Res., 23(226), 2022.*

5 Siehe *Kleinberg et al., Inherent Trade-Offs in the Fair Determination of Risk Scores, LIPIcs, Vol. 67, Schloss Dagstuhl – Leibniz-Zentrum für Informatik (LZI), 2017. DOI: 10.4230/LIPIcs.ITCS.2017.43*

Universität Greifswald
Institut für Mathematik und Informatik
Walther-Rathenau-Straße 47, 17489 Greifswald

Prof. Dr. Alexander Steen
alexander.steen@uni-greifswald.de
Telefon: +49 3834 420 **4634**

Sekretariat: Gaby Aurell
Telefon: +49 3834 420 **4614**
gaby.aurell@uni-greifswald.de

sich an dem Risiko fehlerhafter Ergebnisse orientieren. Dass tatsächlich keine diskriminierenden Algorithmen herausgebildet oder verwendet werden, sollte regelmäßig wissenschaftlich evaluiert werden.

4. Qualität und Authentizität von Daten aus dem Internet

§ 7b AsylG-E ermöglicht es, "öffentlich zugängliche personenbezogene Daten aus dem Internet mit entscheidungserheblichen Angaben aus dem Asylantrag mittels einer automatisierten Anwendung zur Datenverarbeitung [abzugleichen]". Die Begründung gibt dies als Instrument zur Überprüfung angezweifelter Angaben an und erwartet eine effizientere Auswertung öffentlich zugänglicher Informationen (BT-Drs. 21/8275, S. 20–21). Die Aussagekraft eines solchen Instruments hängt jedoch wesentlich von der Qualität der gefundenen Informationen ab. Im Internet öffentlich zugängliche Inhalte können insbesondere sachlich falsch, veraltet oder manipuliert sein oder eine andere Person betreffen. Auch bei einer technisch adäquat durchgeführten Suche können Ergebnisse geliefert werden, die für die Aufklärung des Sachverhalts nicht geeignet sind.

Dabei sind die Authentizität[6], die Integrität[7], die sachliche Richtigkeit und die Zuordnung zur betroffenen Person jeweils individuell zu prüfen: So kann ein tatsächlich von einer Person veröffentlichter Social-Media-Beitrag sachlich falsche Angaben enthalten, ein (inhaltlich) unverändertes Foto mit falschen Metadaten wie Orts- oder Zeitangaben verbreitet werden oder ein scheinbar von einer Person geführtes Benutzerkonto von einem Dritten betrieben werden. Alleine die geklärte Herkunft einer Information ist also noch kein belastbares Argument für die Aussagekraft im Kontext einer konkreten Antragsangabe.[8] Auch die Verbreitung von Fehlinformationen und gezielt irreführenden Inhalten ist zu betrachten, da sich solche Informationen sehr schnell und weit verbreiten können, z.B. in sozialen Netzwerken[9]. Eine hohe Zahl von Fundstellen ist daher allein kein belastbarer Nachweis ihrer sachlichen Korrektheit.

Eine hinreichende Verifizierung kann dabei an fehlenden überprüfbaren Informationen scheitern. Die bloße Untersuchung des gefundenen Inhalts selbst löst dies nicht notwendigerweise auf. Zusätzliche Recherchen können ergänzende Hinweise liefern, die Funde müssen aber ebenfalls auf ihre Qualität und Herkunft geprüft werden. Der erwartete Effizienzgewinn hängt also signifikant davon ab, ob die eingesparte Suchzeit den zusätzlichen Aufwand für Prüfung und Verifizierung in realistischen Einsatzbedingungen übersteigt.

Der Entwurf erkennt mögliche Fehlzuordnungen und unvollständige Datengrundlagen sehr wohl an. § 7b Abs. 1 Satz 3 AsylG-E verlangt dabei jedoch ausschließlich eine manuelle Bewertung und Einordnung als unbedingte Verpflichtung, jedoch nur "gegebenenfalls" eine manuelle Verifizierung. Dabei bleibt weitgehend offen, wann so eine Verifizierung tatsächlich durchgeführt werden muss und welche Anforderungen erfüllt sein müssen, sodass ein gefundener Inhalt zulasten der betroffenen Person verwendet werden darf (BT-Drs. 21/8275, S. 21). Aus informatischer Sicht muss eine tatsächliche Verwendung eine belastbare Prüfung der Quelle, ihrer Authentizität und der konkreten Aussagekraft voraussetzen. Die bloße Auffindbarkeit eines Inhalts im Internet darf nicht zu seiner Würdigung als gesicherte Tatsache führen.

Empfehlung: Der Entwurf sollte verbindliche Anforderungen an die Prüfung und Dokumentation von Internetfunden vorschreiben. Vor einer nachteiligen Verwendung so erlangter Daten sollten Herkunft, Authentizität, Integrität, Zuordnung zur betroffenen Person, zeitlicher und sachlicher Kontext sowie die Unabhängigkeit etwaiger bestätigender Quellen geprüft werden. Zweifel an diesen Voraussetzungen sollten eine Verwendung des Fundes als Grundlage einer nachteiligen Tatsachenfeststellung ausschließen. Der betroffenen Person sollte Gelegenheit gegeben werden, zu den gefundenen Inhalten und ihrer Zuordnung Stellung zu nehmen.

5. Automatisierte Entscheidungsvorlagen

Die Begründung des Entwurfs betont, dass die "rechtliche Würdigung und die abschließende Entscheidung [...] uneingeschränkt bei den zuständigen Mitarbeitenden" verbleiben. Automatisierte Anwendungen sollen weder die individuelle Prüfung ersetzen noch Entscheidungen verbindlich vorprägen (BT-Drs. 21/8275, S. 11). Aus informatischer Sicht muss jedoch zwingend untersucht werden, unter welchen technischen, organisatorischen und prozessualen Bedingungen diese menschliche Kontrolle tatsächlich wirksam wird. Die formale Verantwortung für

6 Die Daten haben tatsächlich die behauptete Herkunft.

7 Die Daten sind gegenüber ihrem Ausgangszustand unverändert und vollständig erhalten geblieben.

8 Siehe *Evaluating Digital Open Source Imagery: A Guide for Judges and Fact-finders, Human Rights Center, UC Berkeley, 2014. URL: https://humanrights.berkeley.edu/publications/evaluating-digital-open-source-imagery-a-guide-for-judges-and-fact-finders/*

9 Siehe *Vosoughi et al., The spread of true and false news online, Science, 359(6380), 2018. DOI: 10.1126/science.aap9559*

Universität Greifswald
Institut für Mathematik und Informatik
Walther-Rathenau-Straße 47, 17489 Greifswald

Prof. Dr. Alexander Steen
alexander.steen@uni-greifswald.de
Telefon: +49 3834 420 **4634**

Sekretariat: Gaby Aurell
Telefon: +49 3834 420 **4614**
gaby.aurell@uni-greifswald.de

eine Entscheidung gewährleistet für sich genommen noch keine eigenständige menschliche Auseinandersetzung mit dem Einzelfall.

Ein erhebliches Risiko ist der sogenannte "Automation Bias". Damit wird die Tendenz beschrieben, automatisiert bereitgestellten Hinweisen oder Empfehlungen ein übermäßiges Vertrauen zukommen zu lassen und diesen ohne hinreichende Prüfung zu folgen. Dies wurde auch bei KI-unterstützten Entscheidungsprozessen beobachtet.[10] Im Kontext der Migrationsverwaltung kann also bereits ein automatisch hervorgehobener vermeintlicher Widerspruch die Sachverhaltsprüfung beeinflussen. Wie stark so ein Einfluss ins Gewicht fällt, hängt von der konkreten Aufgabe und der Ausgestaltung der Arbeitsabläufe ab.

Der Einsatz von Methoden erklärbarer KI *(Explainable Artificial Intelligence, XAI*) gewährleistet hier für sich genommen auch keine hinreichende Kontrolle. Auf diesen Methoden basierende Verfahren sollen zum Beispiel durch die Hervorhebung ausschlaggebender Merkmale nachvollziehbar machen, weshalb ein System eine bestimmte Ausgabe erzeugt hat. Existierende Untersuchungen legen jedoch nahe, dass solche Erklärungen sogar die Wahrscheinlichkeit erhöhen können, einer KI-Empfehlung zu folgen, unabhängig davon, ob diese korrekt ist.[11] Eine verständlich erscheinende Erklärung kann also das Vertrauen in fehlerhafte Empfehlungen verstärken. Die Bereitstellung von Erklärungen sollte daher mit Blick auf ihre tatsächliche Wirkung auf die fachliche Prüfung untersucht werden.

Daraus ergeben sich zusätzliche Anforderungen an die Gestaltung der Verfahren betroffener Behörden: Bei besonders relevanten Bewertungen könnte zum Beispiel geprüft werden, ob die sachbearbeitende Person zuerst eine manuelle Einschätzung dokumentiert, bevor eine automatisierte Ausgabe angezeigt wird. Systemhinweise sollten dabei generell mit den zugrunde liegenden Quellen und erkennbaren Unsicherheiten verknüpft werden. Für entscheidungserhebliche Ausgaben sollte dokumentiert werden, welche darin enthaltenen Informationen tatsächlich geprüft wurden und weshalb die betreffende Ausgabe übernommen oder verworfen wurde. Ein bloßes Bestätigungsfeld kann eine eingehende fachliche Auseinandersetzung nicht nachweisen.

Die Umsetzung solcher Anforderungen muss institutionell abgesichert werden. Sachbearbeitende benötigen ausreichend Zeit, fachliche Qualifikation, Zugriff auf die erforderlichen Daten sowie die Befugnis, Systemausgaben zu verwerfen oder weitere Aufklärung zu veranlassen. Insbesondere sollte eine begründete Abweichung von einer automatisiert generierten Empfehlung organisatorisch ebenso unterstützt werden wie deren begründete Übernahme. Die vorgesehene Schulung nach § 7a Abs. 5 AsylG-E sollte daher praktische Übungen mit plausiblen Fehlhinweisen umfassen (BT-Drs. 21/8275, S. 6-7).

Ob diese Maßnahmen tatsächlich eine menschliche Auseinandersetzung und eigenständige Prüfung unterstützen, sollte vor der Einführung unter realistischen Arbeitsbedingungen untersucht und anschließend im Betrieb laufend überprüft werden. Für die Erprobung können zum Beispiel unabhängig bewertete Fälle mit gezielt eingebrachten Fehlhinweisen zur Verfügung gestellt werden. Maßgeblich ist, ob Beschäftigte diese Fehler erkennen und korrigieren und ob gegenüber einer rein manuellen Bearbeitung die Qualität der Fallbewertung steigt sowie die gesamte Bearbeitungszeit sinkt.

Empfehlung: Der Entwurf sollte die menschliche Entscheidungsverantwortung durch verbindliche Anforderungen an ihre organisatorische und praktische Umsetzung ergänzen. Arbeitsabläufe und Benutzeroberflächen sollten die eigenständige Würdigung des Einzelfalls fördern. Die Wirksamkeit dieser Maßnahmen sollte anhand der Erkennung und Korrektur fehlerhafter Systemhinweise überprüft werden.

6. Datenschutz-Aspekte

Aus technischer Sicht weisen drei datenschutzrelevante Aspekte besonderen Diskussionsbedarf auf.

Bei Internetrecherchen über externe Dienste werden Suchbegriffe sowie je nach technischer Umsetzung Verbindungs- und Nutzungsdaten an die Anbieter der Suchdienste übermittelt. Dazu können IP-Adressen, Zeitpunkte und Identifikatoren gehören.[12] Bereits die Kombination eines Namens mit Angaben wie Herkunft,

10 Siehe *Buçinca et al., To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making, Proc. ACM Hum.-Comput. Interact., 5(CSCW1), 2021. DOI: 10.1145/3449287*

11 Siehe *Bansal et al., Does the Whole Exceed its Parts? The Effect of AI Explanations on Complementary Team Performance, ACM, 2021. DOI: 10.1145/3411764.3445717*

12 Bei der Nutzung von Google als Suchmaschine sind das z.B. Suchbegriffe, IP-Adressen und weitere Nutzungsdaten (siehe Datenschutzerklärung von Google: https://policies.google.com/privacy?hl=de).

Universität Greifswald
Institut für Mathematik und Informatik
Walther-Rathenau-Straße 47, 17489 Greifswald

Prof. Dr. Alexander Steen
alexander.steen@uni-greifswald.de
Telefon: +49 3834 420 **4634**

Sekretariat: Gaby Aurell
Telefon: +49 3834 420 **4614**
gaby.aurell@uni-greifswald.de

Religion oder Verfolgung kann sensible Zusammenhänge offenlegen. Lässt sich die Anfrage zum Beispiel durch IP-Adressenbereiche zudem einer Behörde zuordnen, kann darüber hinaus das Interesse an einer bestimmten Person plausibel dem Aufgabenbereich der Behörde zugerechnet werden. Auch könnten technisch gesehen Suchmaschinen für Suchanfragen von entsprechend erkannten und zugeordneten IP-Adressenbereichen gezielt andere Ergebnisse liefern oder deren Auswahl und Sortierung beeinflussen. Die technische Gestaltung des Internetabgleichs sollte deshalb auch die selbst verursachte Übermittlung von Suchinhalten und Metadaten berücksichtigen und auf das erforderliche Maß begrenzen.

Beim Verfahrensmonitoring scheint insbesondere die Verarbeitung von Klarnamen für die Untersuchung von Bearbeitungszeiten oder wiederkehrenden Verfahrensschritten entbehrlich zu sein. Hierfür kommen auch aggregierte Daten oder Bearbeitungshistorien mit pseudonymisierten Bezeichnern in Betracht. Der Entwurf verlangt zwar die Erforderlichkeit der verarbeiteten Daten, die Begründung führt jedoch die Suche nach bislang unbekannten Strukturen als Argument gegen eine Anonymisierung oder Pseudonymisierung der zu analysierenden Daten an (BT-Drs. 21/8275, S. 17-18). Aus technischer Sicht läuft dieses Argument aber insoweit ins Leere, als die Zuordnung von Verfahrensschritten durch einheitliche Pseudonyme erhalten bleibt. Ist die Auswertung von Zusammenhängen in Verbindung mit konkreten Namen tatsächlich vorgesehen, muss klar dargelegt werden, warum dies für den Untersuchungsgegenstand notwendig ist. Die Verwendung von Klarnamen geht jedoch mit erheblichen Risiken einher, weil sie die Zuordnung sensibler Verfahrensdaten erleichtert. Ein lediglich vermuteter Erkenntnisgewinn ist daher aus technischer Perspektive nicht ausreichend, die zusätzlichen Risiken zu rechtfertigen. Soweit der Zweck einer Anwendung mit anonymisierten Daten erreicht werden kann, ist die zusätzliche Verarbeitung personenbezogener Daten mit dem Grundsatz der Datenminimierung nach Art. 5 Abs. 1 lit. c DSGVO nicht vereinbar. Zu beachten ist, dass pseudonymisierte Daten ebenso personenbezogene Daten darstellen und daher entsprechende Schutzmaßnahmen nötig sind.

Bei der Nutzung personenbezogener Daten zum Training lernender Systeme können weitere Risiken entstehen: Es ist erwiesen, dass in einigen Fällen aus Sprachmodellen einzelne Inhalte aus dem Training, inklusive Namen und weiterer personenbezogener Informationen, durch gezielte Abfragen extrahiert werden können.[13] Das Löschen von Trainingsdaten ist also nicht notwendigerweise ausreichend. Verfahren zum Entfernen des Einflusses bestimmter Trainingsdaten aus bestehenden Modellen (sog. *Machine Unlearning*) sind bekannt, setzen aber spezifische technische Bedingungen voraus und können zusätzlichen Trainingsaufwand nötig machen.[14] Eine zuverlässige, skalierbare und kostengünstige nachträgliche Entfernung bestimmter Daten kann also nicht für beliebige Modelle vorausgesetzt werden.

Empfehlung: (Meta-)Datenübermittlungen an externe Suchdienste sollten vor dem Einsatz erfasst und minimiert werden. Für das Verfahrensmonitoring sollten wann immer möglich anonymisierte Daten verwendet und etwaig nötige Personenbezüge konkret begründet werden. Vor dem Training von statistischen Modellen mit personenbezogenen Daten sollte geklärt sein, wie deren spätere Löschung einschließlich möglicher Auswirkungen auf bereits trainierte Modelle gewährleistet werden kann.

7. Praktische Aspekte

Der Gesetzentwurf verfolgt u.a. das übergeordnete Ziel, die Einheitlichkeit der Verwaltungspraxis zu erhöhen (BT-Drs. 21/8275, S. 1). Die konkrete technische Umsetzung bleibt allerdings zu großen Teilen den beteiligten Behörden überlassen. Insbesondere legt der Entwurf keine gemeinsamen technischen Architekturen oder verbindlichen Schnittstellen fest. Damit entsteht das naheliegende Risiko, dass voneinander unabhängige Anwendungen mit unterschiedlichen Datenformaten, Auswertungsmethoden und Infrastrukturen aufgebaut werden. Ohne gemeinsame fachliche und technische Referenzmaßstäbe können solche Insellösungen bestehende Unterschiede in der Verwaltungspraxis nachhaltig verfestigen oder sogar weitere Unterschiede erzeugen. Dies kann dem erklärten Ziel höherer Einheitlichkeit entgegenwirken.

Bei der Beschaffung von Software und der Nutzung externer Dienstleistungen sind außerdem mögliche Abhängigkeiten von einzelnen Anbietern, Architekturen oder Datenformaten zu berücksichtigen. Proprietäre Datenformate, unzureichend dokumentierte oder unpraktikable Schnittstellen oder ausschließlich beim Anbieter vorhandenes Betriebswissen können einen späteren Wechsel erheblich erschweren und hohe Folgekosten verursachen. Dies berührt insbesondere die digitale Souveränität der beteiligten Behörden. Bereits bei der Beschaffung sollten deshalb offene Standards, ausreichende Prüf- und Nutzungsrechte sowie praktisch

13 Siehe *Carlini et al., Extracting Training Data from Large Language Models, USENIX Association, 2021.*

14 Siehe *Bourtoule et al., Machine Unlearning, IEEE, 2021. DOI: 10.1109/SP40001.2021.00019.*

Universität Greifswald
Institut für Mathematik und Informatik
Walther-Rathenau-Straße 47, 17489 Greifswald

Prof. Dr. Alexander Steen
alexander.steen@uni-greifswald.de
Telefon: +49 3834 420 **4634**

Sekretariat: Gaby Aurell
Telefon: +49 3834 420 **4614**
gaby.aurell@uni-greifswald.de

umsetzbare Möglichkeiten eines Anbieterwechsels berücksichtigt werden. Zudem benötigen die Behörden eigene fachliche und technische Kompetenzen, um die Leistungen externer Anbieter beurteilen und den Betrieb auch bei einer Änderung oder Einstellung des Angebots aufrechterhalten zu können. Soweit personenbezogene Daten an externe Dienstleister übermittelt werden, sind insbesondere die Zwecke, der Umfang und die Bedingungen dieser Verarbeitung unter Berücksichtigung der in Abschnitt 6 beschriebenen Risiken eingehend zu prüfen.

Auch die Darstellung der finanziellen und personellen Auswirkungen ist teilweise fehlleitend. Die Begründung verneint einen Erfüllungsaufwand für die Verwaltung, weil die Einführung automatisierter Anwendungen freiwillig bleibt (BT-Drs. 21/8275, S. 13). Zwar werden erforderliche Investitionen erwähnt, es fehlen aber Angaben, welche Ressourcen tatsächlich nötig sein werden. Die erwarteten Vorteile setzen allerdings voraus, dass geeignete Anwendungen entwickelt oder beschafft, in bestehende Fachverfahren eingebunden und dauerhaft betrieben werden. Es wäre damit nur folgerichtig, entweder die Kosten der Umsetzung zu berücksichtigen oder die erwarteten positiven Auswirkungen stark zu relativieren.

Sollten Entwicklung, Erprobung und Schulung ohne ausreichende zusätzliche Kapazitäten aus dem bestehenden Etat finanziert werden, stehen den Behörden während dieser Zeit gegebenenfalls weniger Mittel und Personal für die eigentliche Fallbearbeitung zur Verfügung. Daraus ergibt sich insbesondere für die Zeit der Einführung ein Risiko längerer Bearbeitungszeiten und höherer Arbeitsbelastung.

Empfehlung: Der Entwurf sollte durch verbindliche und behördenübergreifende Anforderungen an Schnittstellen, Datenbeschreibungen und Qualitätsprüfungen sowie ein Konzept für die Umsetzung erweitert werden, insbesondere unter Berücksichtigung von und Verweis auf bereits existierende Rahmenwerke wie die Verordnung zur Cloud- und KI-Entwicklung (CADA) der EU. Dieses sollte ebenfalls realistische Schätzungen für einmalige und laufende Kosten ausweisen und sicherstellen, dass der Zusatzaufwand während der Einführung solcher Systeme ausfinanziert ist. Unterschiedliche technische Entwürfe und Implementierungen sollten mindestens mittels gemeinsamer Anforderungen und Prüfmaßstäbe vergleichbar sein.

8. Fazit

Automatisierte Anwendungen können zur Verbesserung der Effizienz und Qualität der Migrationsverwaltung beitragen. Die erhofften Vorteile ergeben sich allerdings nicht alleine aus der Befugnis zu ihrer Entwicklung und Nutzung. Es müssen zudem klare Aufgabenzuschnitte, überprüfbare Qualitätsanforderungen, verlässliche Datengrundlagen und eine wirksame menschliche Kontrolle sichergestellt werden. Auch sind einheitliche Standards sowie ausreichende finanzielle und personelle Ressourcen für deren Einführung und Betrieb sowie für die digitale Souveränität der beteiligten Behörden nötig.

Der Gesetzentwurf bleibt hier zu vage und bedarf aus diesem Grund aus technischer Sicht wesentlicher Nachbesserungen. Zulässige Einsatzzwecke, erforderliche Nachweise und organisatorische Bedingungen sollten verbindlich ausgestaltet werden. Andernfalls besteht das Risiko, dass die zukünftig entwickelten Anwendungen bestehende Probleme verstärken oder sogar zusätzliche Belastungen und Fehlerquellen schaffen, ohne die angestrebten Verbesserungen nachweisbar zu erzielen.

In Anbetracht des Umfangs und der Tragweite des Nachbesserungsbedarfs erscheint es angemessen, den Gesetzentwurf grundlegend zu überarbeiten. Dabei sollten die erforderlichen technischen und organisatorischen Konkretisierungen vor einer erneuten parlamentarischen Befassung ausgearbeitet und fachlich geprüft werden.

Universität Greifswald
Institut für Mathematik und Informatik
Walther-Rathenau-Straße 47, 17489 Greifswald

Prof. Dr. Alexander Steen
alexander.steen@uni-greifswald.de
Telefon: +49 3834 420 **4634**

Sekretariat: Gaby Aurell
Telefon: +49 3834 420 **4614**
gaby.aurell@uni-greifswald.de