Konzeptpapier

Ein KI-Transparenzregister für die öffentliche Verwaltung

März 2023

ALGORITHM WATCH

Konzeptpapier

/ EIN KI-TRANSPARENZREGISTER FÜR DIE ÖFFENTLICHE VERWALTUNG

März 2023

Die öffentliche Verwaltung kann Automatisierungs- und KI-Verfahren nutzen, um ihre Angebote und Prozesse erheblich zu verbessern. Um gleichzeitig die Digitalkompetenzen der Behörden zu fördern, die Rechte der Bürger*innen zu schützen und das Gemeinwohl zu stärken, sollten Bund und Länder ein KI-Transparenzregister einrichten. Die KI-Verordnung der EU wird nicht ausreichen, um Softwaresysteme der Behörden angemessen zu prüfen und transparent zu machen. Über ein Online-Register sollten die wichtigsten Informationen über KI-Verfahren, die in der öffentlichen Verwaltung zum Einsatz kommen, öffentlich abrufbar sein. Der Veröffentlichung geht ein strukturiertes Verfahren voraus, das die Behörden intensiv bei der Konzeption und Entwicklung ihrer algorithmischen Systeme unterstützt. Die Potenziale der Automatisierung lassen sich nur dann nutzen, wenn die Behörden eine fundierte Einschätzung zu der Frage erlangen, wie die Systeme zielführend eingesetzt werden können und wie sich unerwünschte Folgen vermeiden lassen.

Die aktuelle Bundesregierung hat sich als zentrales Projekt vorgenommen, die staatlichen Strukturen und Arbeitsprozesse zu modernisieren. Die Verwaltung soll zeitgemäß, digital und bürger*innenorientiert ausgebaut werden. Anträge stellen, Termine buchen, Leistungen beantragen: All das kostet uns beim Gang zum Amt viel Zeit. Digitalisierung und Automatisierung leisten genauso wie der Einsatz von Künstlicher Intelligenz (KI) bzw. von automatisierten Entscheidungssystemen (automated decision making, ADM) einen wichtigen Beitrag dazu, dass der Staat den Menschen einen guten Service bietet und mit den Menschen zeitgemäß kommuniziert.

Behörden tragen eine besondere Verantwortung gegenüber denen, die von ihren Entscheidungen betroffen sind. Daher müssen automatisierte Prozesse verantwortungsvoll, grundrechtskonform und gemeinwohlorientiert zum Einsatz kommen. Das bedeutet beispielsweise, dass automatisierte Systeme, die über die Vergabe von Sozialleistungen oder Schulplätzen entscheiden, nicht zu Diskriminierung oder intransparenten Entscheidungen führen dürfen.

Auf europäischer Ebene soll die KI-Verordnung (auch *Artificial Intelligence Act*, AI Act) festlegen, welchen Anforderungen KI-Systeme genügen müssen – abhängig von dem Risiko, das sie für Mensch, Nutzer*innen und Gesellschaft darstellen. In einer EU-Datenbank sollen Hersteller und Betreiber KI-Systeme mit hohem Risiko registrieren – ohne dabei aber im Einzeln erläutern zu müssen, wie ihre KI funktioniert und sie eingesetzt wird. Allerdings trägt häufig der Kontext, in dem ein System eingesetzt wird, wesentlich zu den Risiken bei, die von ihm ausgehen.

Aber nicht nur im Sinne der Risikovermeidung ist es problematisch, wenn die Einsatzszenarien von KI-Systemen im Dunkeln bleiben: Wenn unklar ist, welche ADM-Systeme die Verwaltung wo einsetzt, hat das sowohl Nachteile für Verwaltungen selbst, die nicht aus den Erfahrungen anderer lernen können, als auch für Unternehmen, die daran beteiligt sein möchten und sollen, Lösungen für Behörden zu entwickeln.

Ein KI-Transparenzregister für die öffentliche Verwaltung bietet viele Vorteile, die sich auf verschiedene gesellschaftliche Akteure positiv auswirken würden:

— Transparenz über den Einsatz von ADM-Systemen würde den Verwaltungen helfen, aus den Hürden und Erfolgen anderer Projekte zu lernen. Zu sehen, wie andere öffentliche Akteure tech-

nische Systeme einsetzen, um ihre Aufgaben zu erfüllen, würde einen Austausch zwischen Behörden ermöglichen, um eine gute digitale Verwaltung aktiv voranzutreiben.

— Innovationsorientierte Unternehmen und Start-ups würden einen Überblick über ADM-Systeme erhalten, die bereits zum Einsatz kommen. Sie könnten so bessere *GovTech* entwickeln und anbieten, die den Bedürfnissen der Behörden entsprechen.

— Menschen könnten automatisierte Entscheidungen, von denen sie betroffen sind, besser nachvollziehen und von Schutzrechten Gebrauch machen.

— Zivilgesellschaft und Wissenschaft würden einen Überblick darüber bekommen, welche Automatisierungssysteme in der deutschen Verwaltung im Einsatz sind. Es würde ein lebendiger Dialog darüber entstehen, welche Innovationen und welches gesellschaftliche Zusammenleben wir uns wünschen.

/ Wie sollte das KI-Register aufgebaut sein?

— Die Informationen im KI-Transparenzregister sollten öffentlich verfügbar sein, beispielsweise auf dem GovData-Portal.

— Das Register sollte alle relevanten Daten über den Einsatz von Automatisierungssystemen in der öffentlichen Verwaltung enthalten. Zu diesen relevanten Daten gehören mindestens Informationen zur eingesetzten Software, zum Zweck des Systems, zu den an Entwicklung und Einsatz beteiligten Akteuren, grundlegende Informationen zum Entscheidungsmodell und der Architektur, zu verwendeten Daten sowie die Ergebnisse einer Folgenabschätzung (s. dazu den Abschnitt unten).

— Die im Register enthaltenen Informationen sollten für die Öffentlichkeit verständlich, aktiv, vollständig, unverzüglich, kosten- und barrierefrei, nicht proprietär und lizenzfrei in einem maschinenlesbaren und interoperablen Format bereitgestellt werden, d.h. sie müssen nach Vorgabe eines standardisierten Protokolls strukturiert sein.

— Einheitliche Standards und offene Schnittstellen sind notwendig, um das Register mit anderen deutschen oder europäischen Registern zu verknüpfen und um einen Zugang für wissenschaftliche und zivilgesellschaftliche Akteure zu schaffen.

Zivilgesellschaft und Wissenschaft sollten die Entwicklung und den Aufbau des KI-Transparenzregisters im Rahmen eines Beteiligungsprozesses aktiv mitgestalten und fortlaufend evaluieren.

/ Welche Informationen sollten im KI-Transparenzregister einsehbar sein?

1. **Zweck und konkrete Einbettung des Systems im jeweiligen Entscheidungsprozess** – Beispielfragen, die zu beantworten sind: Für welches Problem soll das System eine Lösung liefern? Wird das System verwendet, um Entscheidungen über Personen zu treffen, Empfehlungen zu geben oder Entscheidungen zu beeinflussen? Mit welchen Daten und Entscheidungsmustern arbeitet das System? Findet eine menschliche Kontrolle statt?

2. **Akteure, die an Entwicklung und Einsatz eines ADM-Systems beteiligt sind** – Beispielfragen, die zu beantworten sind: Wer ist für die Konzeption, Entwicklung und Implementierung des Systems verantwortlich? Wer ist für den konkreten Einsatz des Systems und dessen Resultate verantwortlich? Wer ist verantwortlich für die Verwaltung der Antworten und Rückmeldungen der Endnutzer*innen, d. h. der Personen, die das System benutzen oder von ihm unterstützt werden? Wer ist dafür verantwortlich, auf Zweifel oder Herausforderungen von Personen zu antworten, die von der Nutzung des Systems betroffen sind? Welche Stakeholder hat die zuständige Stelle zu welchem Zweck an welcher Stelle des

Einführungsprozesses wie beteiligt? Welche Auswirkung hatte ihre Beteiligung auf die Konzeption und Einführung des Systems?

3. **Grundlegende Informationen zum Entscheidungsmodell des ADM-Systems** – Beispielfragen, die zu beantworten sind: Beruht das technische System auf einem statistischen Modell menschlichen Verhaltens oder auf einem Modell einzelner persönlicher Merkmale? Ist das System so konzipiert, dass es adaptiv ist (es also seine Parameter ändert, z. B. um effizienter zu werden), so dass nicht alle neuen Fälle wie frühere behandelt werden?

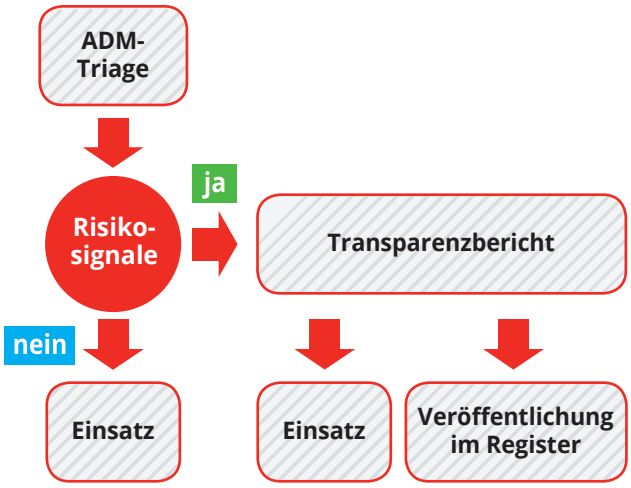
4. **Informationen zu den verwendeten Methoden** – Beispielfragen, die zu beantworten sind: Welche Methoden kamen zur Anwendung, um die Voreingenommenheit und die Fairness des Systems zu definieren und zu messen? Mit welchen Methoden wurde die Leistung des Systems getestet und gemessen? Mit welchen Methoden wurden die von den Systemvorhersagen/-empfehlungen/-entscheidungen unmittelbar betroffenen Stakeholder identifiziert – und was sind die voraussichtlichen Auswirkungen auf diese Personen?

5. **Qualitätssicherungsprozesse, Maßnahmen zur Informationssicherheit und zum Datenschutz** – Beispielfragen, die zu beantworten sind: Welche Prozesse der Qualitätssicherung wurden dem Einsatz vorangestellt, welche kommen fortlaufend zum Einsatz? Welche Maßnahmen zur Informationssicherheit und zum Datenschutz sind vorgesehen, beispielsweise durch IT-Abteilungen der Behörden und Datenschutzbeauftragte?

/ Wie werden diese Informationen erhoben?

Bei der Einführung eines automatisierten Systems sollten rechtliche, ethische und gesellschaftliche Anforderungen von Beginn an reflektiert und in das System integriert werden. Zu diesem Zweck sollte die einführende Behörde selbstständig mittels einer Folgenabschätzung strukturiert und standardisiert relevante Informationen erheben, die mindestens im KI-Transparenzregister abgebildet sein sollten. Diese Folgenabschätzung würde zugleich dazu führen, dass die Verwaltung interne Kompetenzen über die rechtlichen, ethischen und gesellschaftlichen Anforderungen aufbaut. Dabei könnte sie Unterstützung von einer *Taskforce KI-Kompetenz* erhalten: Diese Taskforce würde Behörden bei der Einführung mit Erfahrungswissen aus anderen Projekten unterstützen und beim Aufbau neuer Kompetenzen begleiten.

AlgorithmWatch hat gemeinsam mit dem Kanton Zürich und der Universität Basel eine Folgenabschätzung für automatisierte Systeme erarbeitet, die sich dafür als Grundlage nutzen lässt. Mittels geeigneter Verfahren sollte sie in einen formalen Leitfaden für die Verwaltung überführt werden.



/ Wie funktioniert die Folgenabschätzung?

Das Instrument unterscheidet sich von anderen Vorschlägen durch seine zwei Stufen. Auf der ersten Stufe wird in einem möglichst einfachen Verfahren ermittelt, ob vom ADM-System überhaupt Risiken ausgehen. Dieser Schritt reagiert darauf, dass sich „Künstliche Intelligenz“ nicht klar definieren lässt und man nicht annehmen kann, dass alle KI-basierten Systeme risikobehaftet sind – während zugleich Risiken auch von Systemen ausgehen können, die keine Verfahren verwenden, die auf KI-Technologien fußen. So ist z.B. ist eine automatische, auf Methoden des Maschinellen Lernens basierende Rechtschreibprüfung Teil

eines sozio-technischen Systems. Wenn Menschen, die Entscheidungen über Individuen treffen, das System verwenden, ließe es sich als „verwendet, um Entscheidungen über Individuen zu treffen" beschreiben. Aber es beeinflusst in keiner (erkennbaren und wissenschaftlich plausiblen) Weise, wie die Entscheidung zustande kommt. In diesem Verfahren würde also kein Risikosignal davon ausgehen, dass eine automatisierte Prüfung zur Anwendung kommt. Dagegen kann ein System, das automatisiert Kindergeld auszahlt, große Risiken bergen, auch wenn es auf einem Verfahren beruht, das nicht als „KI" angesehen wird.

/ So wenig Bürokratie wie möglich, so viel Transparenz wie nötig

Diese erste Stufe der Folgenabschätzung – die sogenannte Triage von ADM-Systemen – vermeidet unnötige Bürokratie: Wenn keine Risikosignale auftauchen, kann das System sofort eingesetzt und es muss kein Transparenzbericht erstellt werden. Bei wenigen Signalen sind die Anforderungen an den Bericht gering. Mit Hilfe der Folgenabschätzung lässt sich also auch nach unten hin abgrenzen, wie aufwendig der jeweilige Eintrag in das KI-Transparenzregister zu sein hat. Zugleich wird im Zuge dieses Verfahrens angemessen reflektiert, ob das ADM-System gut genug konzipiert ist, um es umzusetzen. Die Folgenabschätzung ist auch während des Einsatzes eines Systems fortlaufend durchzuführen. Aus den Ergebnissen der Folgenabschätzung entsteht letztlich der Bericht, der in das öffentliche Transparenzregister einfließt.

Nur in Ausnahmefällen – etwa, wenn nicht alle Informationen aus legitimen Geheimhaltungsinteressen (wie dem Schutz der Privatsphäre) vollständig offengelegt werden dürfen – ist dies im Register zu vermerken und die Aufsichtsbehörde anzugeben, der gegenüber die Angaben vollständig offengelegt wurden.

/ Wie wird das KI-Transparenzregister überprüft?

Das KI-Transparenzregister kann seine Funktion nur dann erfüllen, wenn es verlässlich und vollständig ist und eine öffentliche Kontrolle ermöglicht. Behörden sollten verpflichtet sein, alle relevanten Informationen in dem Register anzugeben. Bei fehlenden oder falschen Angaben sollten Sanktionen erfolgen. Die Angaben werden von einer Aufsichtsbehörde stichprobenartig überprüft.

Durch die Kombination aus KI-Register und Folgenabschätzung, bei der die Behörden intensive Unterstützung durch die Digitalexpertise einer Taskforce KI-Kompetenz erhalten, kann es gelingen, die Vorteile von Automatisierungsverfahren für Verwaltung und Bürger*innen optimal zu nutzen. Der moderne Staat agiert mit offenem Visier, kooperiert mit Wissenschaft und Zivilgesellschaft und setzt digitale Techniken nur dann ein, wenn sie dem Gemeinwohl dienen und keine unangemessenen Risiken herbeiführen.

LINKS

Sie können dieses Konzeptapier als PDF hier herunterladen:

https://algorithmwatch.org/de/transparenzregister-oeffentliche-verwaltung-2023/

Und hier finden Sie unser Instrument zur Folgenabschätzung (Impact Assessment Tool) zur Analyse automatisierter Entscheidungssysteme im öffentlichen Sektor:

https://algorithmwatch.org/de/adms-impact-assessment-public-sector-algorithmwatch/

KONTAKT

Pia Sombetzki
Policy & Advocacy Managerin
AlgorithmWatch
sombetzki@algorithmwatch.org
030 - 99 40 49 006

Matthias Spielkamp
Geschäftsführer AlgorithmWatch
spielkamp@algorithmwatch.org