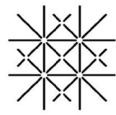




Universität
Zürich ^{UZH}



Universität
Basel

Juristische
Fakultät

WHITE PAPER

Transparenz durch öffentliches Verzeichnis

Nadja Braun Binder
Liliane Obrecht

Juni 2024

CENTER FOR
INFORMATION
TECHNOLOGY
SOCIETY AND
LAW — ITS_L

e-PIAF

electronic Public Institutions and
Administrations Research Forum

Dieses White Paper wurde im Projekt «**Nachvollziehbare Algorithmen: ein Rechtsrahmen für den Einsatz von Künstlicher Intelligenz**» entwickelt, das vom Center for Information Technology, Society, and Law (ITSL) der Universität Zürich und von der Forschungsstelle electronic Public Institutions and Administrations Research Forum (e-PIAF) der Universität Basel durchgeführt und von der Stiftung Mercator finanziell unterstützt wird. Dieses White Paper ist Teil einer Reihe von White Papers, die sich mit den zentralen Herausforderungen befassen, die mit dem Einsatz von Künstlicher Intelligenz (KI) in Unternehmen und in der Verwaltung verbunden sind.

Die White Papers und weitere Informationen zum Projekt finden sich auf www.itsl.uzh.ch und www.ius.unibas.ch/e-piaf.

Folgende White Papers sind bislang erschienen:

- **Manipulation**
- **Diskriminierung**
- **Datenschutz**
- **Transparenz**
- **Transparenz durch Begründung von Verfügungen**
- **Transparenz durch öffentliches Verzeichnis**

Das Kernprojektteam besteht aus folgenden Personen:

Prof. Dr. Florent Thouvenin, Professor für Informations- und Kommunikationsrecht an der Universität Zürich, Vorsitzender des Lenkungsausschusses des ITSL

Prof. Dr. Nadja Braun Binder, MBA, Professorin für öffentliches Recht an der Universität Basel

Dr. Stephanie Volz, wissenschaftliche Geschäftsführerin des ITSL und Lehrbeauftragte an der Universität Zürich

Liliane Obrecht, MLaw, Wissenschaftliche Mitarbeiterin und Doktorandin an der Juristischen Fakultät der Universität Basel

Der Einsatz algorithmischer Systeme kann dazu führen, dass staatliche Entscheidungen nicht mehr ohne Weiteres nachvollziehbar sind. Zur Erhöhung des Vertrauens der Bevölkerung in solche Entscheidungen und zur Sicherstellung von externen Kontrollmöglichkeiten sollte allgemein ersichtlich sein, wann ein algorithmisches System für bestimmte Entscheidungen eingesetzt wird. Mit einem öffentlichen Verzeichnis über den Einsatz algorithmischer Systeme kann die notwendige Transparenz hergestellt werden.

Problemstellung

In jüngster Zeit hat das Thema der Herstellung von Transparenz über den Algorithmen Einsatz in öffentlichen Verwaltungen erhöhte Aufmerksamkeit erfahren. So wurden Ende 2021 und im Mai 2023 im Nationalrat zwei Vorstösse eingereicht, mit welchen der Bundesrat aufgefordert wird, ein öffentliches Verzeichnis aller in der Bundesverwaltung eingesetzten algorithmenbasierten, automatisierten Entscheidungssysteme anzufertigen (Motion 21.4508 Schlatter; Postulat 23.3566 Schlatter). Auch der Zürcher Kantonsrat hat im April 2022 ein Postulat überwiesen, in welchem der Regierungsrat gebeten wird, ein Verzeichnis zur Erfassung der eingesetzten automatisierten Entscheidungssysteme zu erstellen (Postulat 9/2022 Yuste/Moser/Mäder/Willi/Zahler). Mittlerweile liegt ein Gesetzesentwurf im kantonalen Informations- und Datenschutzgesetz (IDG-ZH) vor, der als Grundlage für die Pflicht zur Führung eines solchen Verzeichnisses dient (§ 13 Abs. 3 revidiertes IDG-ZH). Im Kanton Basel-Stadt wurde im Oktober 2023 ein Anzug betreffend Transparenz über den Einsatz von algorithmischen Systemen in der Verwaltung eingereicht (Anzug 23.5474 der Justiz-, Sicherheits- und Sportkommission). In allen genannten parlamentarischen Vorstössen geht es um algorithmische Systeme, die unter dem Gesichtspunkt der Nachvollziehbarkeit der Entscheidungsfindung problematisch erscheinen.

Die Herstellung von Transparenz über den staatlichen Algorithmen Einsatz stellt ein Instrument zur Sicherstellung einer **rechtsstaatskonformen** und **diskriminierungsfreien digitalen Transformation der öffentlichen Verwaltung** dar. Es ist oft unklar, wo algorithmenbasierte Systeme überhaupt eingesetzt werden. Da der Staat seiner Bevölkerung in der Regel hoheitlich gegenübertritt, bedarf es gewisser Mechanismen, um staatliches Handeln kontrollierbar zu machen. Zusätzlich zu allenfalls existierenden bzw. noch

zu schaffenden individuellen Kontrollmöglichkeiten im Einzelfall für unmittelbar Betroffene, soll insbesondere die Zivilgesellschaft die Möglichkeit erhalten, den staatlichen Algorithmen Einsatz kritisch zu begleiten.

Dadurch kann die **Vertrauensbeziehung** zwischen dem Staat und der Bevölkerung auch bei zunehmender Nutzung digitaler Technologien erhalten und sogar gestärkt werden. Transparenz kann die Vertrauenswürdigkeit fördern, indem die Bevölkerung die Möglichkeit erhält, sich möglichst umfassend zu informieren. Dadurch wird eine zusätzliche Motivation für **rationales Staatshandeln** sowie **behördliche Selbstkontrolle** geschaffen.

Transparenz über den Einsatz algorithmischer Systeme kann ausserdem dabei helfen, allfällige **Diskriminierungsprobleme** sichtbar zu machen, indem durch Vergleich verschiedener mit demselben System bearbeiteter Fälle erkennbar wird, ob eine Entscheidung an ein verpöntes Merkmal (siehe Art. 8 Abs. 2 BV, z. B. Herkunft, Geschlecht) anknüpft.

Algorithmische Systeme und Künstliche Intelligenz

Im vorliegenden Kontext liegt der Fokus auf **algorithmischen Systemen**, die aufgrund erswerter Nachvollziehbarkeit problembehaftet sein können. Dazu gehören Systeme, die **ohne Mitwirkung einer natürlichen Person** funktionieren und z. B. zum (vollautomatisierten) Erlass sogenannter automatisierter Einzelentscheidungen genutzt werden. Auch **KI-Systeme** zählen dazu, die insbesondere aufgrund der erswerter Nachvollziehbarkeit des Zustandekommens ihrer Outputs eine Herausforderung darstellen können. Bilden die Ausgaben von KI-Systemen eine Grundlage für menschliche Entscheidungen

(Entscheidungsunterstützung), die sich direkt oder indirekt auf die Rechte und Pflichten natürlicher Personen auswirken, oder handelt es sich um ein algorithmisches System, das selbst eine Entscheidung trifft, die sich direkt oder indirekt auf die Rechte und Pflichten von Personen auswirkt, so können sie als Systeme zur algorithmischen Entscheidungsfindung («Algorithmic Decision Making System», ADMS) zusammengefasst werden. Im Folgenden wird daher von einem **ADMS-Verzeichnis** gesprochen.

Transparenzbegriff

Zum Begriff der Transparenz findet sich im Bundesrecht **keine Legaldefinition**. Es existiert auch keine andere einheitliche Definition des Begriffs. Dennoch durchzieht der Transparenzbegriff die gesamte **Bundesverfassung** (BV, SR 101) und kommt in unterschiedlichen Bestimmungen zum Ausdruck. So kann er beispielsweise als Element der **Rechtsstaatlichkeit** (Art. 5 BV) verstanden oder als Ausfluss grundrechtlicher **Verfahrensgarantien** (siehe White Paper «Transparenz durch Begründung von Verfügungen»), eingeordnet werden. Ein weiterer verfassungsrechtlicher Anknüpfungspunkt stellt der **Informationsauftrag** des Bundesrates nach Art. 180 Abs. 2 BV dar. Gemeinsam ist diesen Anknüpfungspunkten, dass sie insbesondere der Stärkung der Vertrauensbeziehung zwischen Staat und Bevölkerung sowie der Legitimität staatlichen Handelns dienen.

Auf **Gesetzesstufe** wird Transparenz häufig im Rahmen des Öffentlichkeitsprinzips der öffentlichen Verwaltung diskutiert. Zentrales Anliegen des seit 2004 auf Bundesebene geltenden **Öffentlichkeitsgesetzes** (BGÖ, SR 152.3) ist die Verbesserung der Kontrolle der Bevölkerung über die Verwaltung, wodurch das Vertrauen in diese gestärkt werden soll. Das BGÖ fördert die Transparenz der Verwaltung, indem jede interessierte Person in die Dokumente der Bundesbehörden Einsicht nehmen darf. Allerdings erfasst das BGÖ lediglich die passive Informationstätigkeit der Verwaltung, eine aktive Information durch Behörden ist nicht vorgesehen.

Die aktive Informationstätigkeit findet ihre Grundlage insbesondere in Art. 10 des **Regierungs- und Verwaltungsorganisationsgesetzes** (RVOG, SR 170.010), der den Informationsauftrag gemäss Art. 180 Abs. 2 BV konkretisiert. Der Informationspflicht unterliegen die Regierung sowie die Bundesverwaltung, wobei Umfang und Inhalt der Information vom jeweiligen Ort und der

jeweiligen Zeit abhängig sind. Ausschlaggebend sind die Bedürfnisse der Bevölkerung sowie des Staates selbst.

Auch das **Datenschutzgesetz** (DSG, SR 235.1) enthält verschiedene Vorgaben zur Herstellung von Transparenz. Dabei geht es um die Herstellung von Transparenz bei der Bearbeitung von Personendaten. Diesem Zweck dienen etwa der Grundsatz, dass die Beschaffung von Personendaten und der Zweck, zu dem sie beschafft werden, erkennbar sein müssen (Art. 6 Abs. 3 DSG), die Pflicht der Verantwortlichen, die betroffene Person über die Datenbeschaffung zu informieren (Art. 19 DSG), die Kennzeichnungspflicht bei einem automatisierten Einzelentscheid durch ein Bundesorgan (Art. 21 Abs. 4 DSG) oder das Verzeichnis der Datenbearbeitungstätigkeiten von Bundesorganen (Art. 56 DSG).

Lösungsansatz ADMS-Verzeichnis

Von verschiedenen Seiten wird die Schaffung eines öffentlich zugänglichen ADMS-Verzeichnisses der in der öffentlichen Verwaltung eingesetzten algorithmischen Systeme gefordert, um den Transparenzanforderungen nachzukommen und gesellschaftliche Kontrolle zu konkretisieren. Diese Überlegungen sind auch aus rechtswissenschaftlicher Perspektive sinnvoll.

Das Führen öffentlicher Verzeichnisse ist aus rechtlicher Sicht nichts Neues. Seit jeher werden öffentliche Verzeichnisse in der Rechtswissenschaft als **Mittel zur Transparenzherstellung** und somit zur Sicherstellung von Rechtsstaatlichkeit betrachtet. Ein Verzeichnis ist ein geordnetes Informationssystem, das Orientierung über einen bestimmten Sachbereich bietet. Es ist öffentlich, sobald jede Person zu einem unspezifischen Zeitpunkt und ohne spezifisches Interesse in dieses Einsicht nehmen kann.

Ein ADMS-Verzeichnis verfolgt mehrere Ziele. Zentral ist die **Pflege des Vertrauensverhältnisses** zwischen dem Staat und der Bevölkerung. Da der Staat hoheitlich, d.h. einseitig, über Rechte und Pflichten der Bevölkerung entscheiden kann, entsteht ein Machtungleichgewicht, das potenziell zu Machtmissbrauch führen kann. Indem der Staat jedoch offenlegt, wo er welche ADMS einsetzt, schafft er Kontrollmechanismen und fördert die Akzeptanz solcher Technologien in der Bevölkerung. Darüber hinaus kann ein Verzeichnis auch einer möglichen Intransparenz über die Funktionsweise von ADMS entgegenwirken, indem es

diese näher erläutert. Von besonderer Bedeutung ist dabei die Schaffung einer **adressatengerechten Transparenz**. So sollten neben leicht verständlichen Informationen für Laien auch weiterführende Dokumentationen für Expertinnen und Experten zur Verfügung gestellt werden. Schliesslich wird die Verwaltung selbst durch die Dokumentation entlastet, da sich Interessierte direkt über die eingesetzten Systeme informieren können und kein einzelfallweiser verwaltungsinterner Aufwand für die Auskunftserteilung im Einzelfall anfällt.

Rechtliche Erfassung

Behörden dürfen aufgrund ihres Informationsauftrages – wie er z. B. auf Bundesebene in Art. 10 RVOG verankert ist – bereits heute ein ADMS-Verzeichnis führen. Die Nachteile von solchen freiwillig geführten ADMS-Verzeichnissen sind, dass sie keinem einheitlichen Aufbau folgen und auch nicht an zentraler Stelle zugänglich gemacht werden müssen. Darüber hinaus ist bei einer rein freiwilligen Verzeichnisführung nicht klar, ob auch wirklich alle Systeme erfasst werden. Deshalb ist die **Pflicht zur Führung eines ADMS-Verzeichnisses auf Gesetzesstufe** zu verankern und in den Grundzügen festzulegen, welche Angaben ein solches Verzeichnis enthalten muss, wer das Verzeichnis zu verwalten und wer die Angaben einzutragen hat.

Für die Bundesebene kämen verschiedene bereits bestehende Gesetze als Grundlage infrage. Dazu zählen das DSG oder das am 1. Januar 2024 in Kraft getretene Bundesgesetz über den Einsatz elektronischer Mittel zur Erfüllung von Behördenaufgaben (EMBAG, SR 172.019). Grundsätzlich wäre auch das RVOG ein möglicher Regelungsort. Das EMBAG ist jedoch spezifisch auf Fragen der Digitalisierung ausgerichtet und weist daher einen engeren Sachbezug für die Regelung eines ADMS-Verzeichnisses auf als das RVOG.

Entscheidend für die Frage der **rechtlichen Verankerung** eines ADMS-Verzeichnisses ist neben dem Sachbezug letztlich der Adressatenkreis. Während das **DSG** alle Bundesorgane erfasst, die eine öffentliche Aufgabe wahrnehmen (Art. 2 Abs. 1 lit. b i. V. m. Art. 5 lit. i DSG), weist das **EMBAG** einen eingeschränkten Adressatenkreis auf. Es richtet sich nur an die zentrale Bundesverwaltung. Zwar kann der Bundesrat auf Verordnungsstufe eine Ausweitung des Adressatenkreises vorsehen (Art. 2 Abs. 1, 2 und 3 EMBAG), doch sind damit in erster Linie punktuelle Ausweitungen gemeint und keine generelle Ausweitung auf alle Bundesorgane, auch wenn dies nur die Vorgaben für ein

ADMS-Verzeichnis betreffen würde. Ziel eines ADMS-Verzeichnisses ist es, möglichst alle in der öffentlichen Verwaltung eingesetzten ADMS zu erfassen. **Ziel führend** erscheint daher die **Verankerung im DSG**. Ansprüche nach dem BGÖ würden durch die Regelung des ADMS-Verzeichnisses nicht berührt.

Konkretisierungen zu den einzelnen Elementen des Verzeichnisses können schliesslich auf **Verordnungsebene** vorgenommen werden und müssen nicht auf Gesetzesstufe verankert werden. Dazu zählen unter anderem die folgenden Aspekte:

– Basisinformationen

Als grundlegende Information dient erstens eine allgemeine **Übersicht**, die allen Personen in verständlicher Sprache Erklärungen über die Funktionsweise des algorithmischen Systems, die Art und Weise seines Einsatzes und dessen Auswirkungen liefert. Zweitens sind Informationen zum **verantwortlichen Organ** aufzuführen. Dieses stellt sicher, dass das System über seinen gesamten Lebenszyklus weiterentwickelt und gewartet wird. In diesem Rahmen können auch die an der Entwicklung des Systems beteiligten Personen sowie diejenigen, die das System betreiben, genannt werden. Wichtig ist jedoch, dass klar ersichtlich wird, wer die Verantwortung trägt. Die genannten Personen müssen kontaktiert werden können (beispielsweise via E-Mail oder Telefon). Drittens sind gegebenenfalls weitere **Referenzen** zu nennen, die weiterführende Informationen (auch in Bezug auf Details technischer Natur) und allfällige Prüfungsberichte liefern können.

– Einsatzbereich des ADMS

Das Verzeichnis muss Informationen über den Bereich, in dem das ADMS eingesetzt wird, enthalten. Dazu gehört auch die **Rechtsgrundlage**, die den Einsatz eines Systems legitimiert.

– Daten

Der **Zweck** der Bearbeitung von Personendaten ist anzugeben. Falls der Zweck aus der gesetzlichen Grundlage hinreichend hervorgeht, kann auf diese verwiesen werden. Ebenso soll das Verzeichnis Auskunft zur Art und Herkunft der bearbeiteten Sach- und Personendaten (**Datensets**) geben. Dabei müssen keine Einzeldaten veröffentlicht, sondern die einschlägigen Merkmale oder Kategorien genannt werden. Allfällige Geheimhaltungsinteressen sowie die Anonymisierung von Personendaten sind zu

berücksichtigen. Auch die Art und Weise, wie das algorithmische System die Datenverarbeitung durchführt (**Funktionsweise**) und wie die Resultate erzielt werden, bedarf der Konkretisierung auf Verordnungsstufe. Die Bearbeitungsmittel und die Logik des Algorithmus sind aufzuzeigen. Zudem müssen Informationen zur Interpretation und Messbarkeit des Systems aufgeführt werden. Dies kann etwa durch das Festhalten der Hauptmerkmale realisiert werden. Die Offenlegung des Quellcodes hingegen ist nicht zwingend erforderlich. Dieser könnte allerdings ausgewählten Expertinnen und Experten offenbart bzw. bei Bedarf bereitgestellt werden. Ist ein System im Einsatz, das mithilfe von **Trainings-, Validierungs- oder Testdaten** entwickelt wurde, muss ersichtlich sein, welche Daten für das Training, die Validierung bzw. das Testen genutzt wurden.

– **Algorithmus**

Ferner sind Informationen zu den eingesetzten Algorithmen notwendig, insbesondere darüber, ob es sich um **regelbasierte** oder **datenbasierte** («Machine Learning») Algorithmen oder auch um eine Kombination aus beiden handelt.

– **Auswirkungen auf die Bevölkerung**

Im öffentlichen Verzeichnis sollten Methoden be-

schrieben werden, die bei der Implementierung und beim Einsatz des algorithmischen Systems angewendet werden, um potenzielle Diskriminierungsrisiken zu adressieren und zu verhindern (**Antidiskriminierung**). Zusätzlich bedarf es bei der Offenlegung der Resultate einer **Folgenabschätzung**, in der die Risiken des jeweiligen algorithmischen Systems genannt werden. Dies bildet die Grundlage für die Festlegung geeigneter Massnahmen zum Schutz der Bevölkerung vor bzw. zur Bewältigung der erkannten Risiken.

– **Menschliche Aufsicht**

Die Behörde muss in der Lage sein, die eingesetzten Systeme durch Menschen zu beaufsichtigen. Die Ausgestaltung der menschlichen Aufsicht hängt vom Grad der Autonomie des algorithmischen Systems ab. Grundsätzlich sollte aus dem ADMS-Verzeichnis hervorgehen, inwieweit eine menschliche Aufsicht möglich und vorhanden ist. Darüber hinaus sind die für die menschliche Aufsicht erforderlichen Fähigkeiten zu beschreiben.

– **Metadaten**

Schliesslich müssen Informationen zu den Metadaten angegeben werden, z. B. wann ein ADMS zum ersten Mal in das Verzeichnis aufgenommen wurde oder wann Informationen überarbeitet wurden.

Impressum

© 2024

Herausgeberin:
Center for Information Technology,
Society, and Law (ITSL)
Universität Zürich
Rämistrasse 74|38
8001 Zürich