Sachdokumentation:

Signatur: DS 3728

Permalink: www.sachdokumentation.ch/bestand/ds/3728

Zitierweise für graue Literatur

Elektronische Broschüren und Flugschriften (DS) aus den Dossiers der Sachdokumentation des Sozialarchivs werden gemäss den üblichen Zitierrichtlinien für wissenschaftliche Literatur wenn möglich einzeln zitiert. Es ist jedoch sinnvoll, die verwendeten thematischen Dossiers ebenfalls zu zitieren. Anzugeben sind demnach die Signatur des einzelnen Dokuments sowie das zugehörige Dossier.

University of Zurich[UZH]

Digital Society Initiative

Positionspapier

Ein Rechtsrahmen für Künstliche Intelligenz

Florent Thouvenin, Markus Christen, Abraham Bernstein, Nadja Braun Binder, Thomas Burri, Karsten Donnay, Lena Jäger, Mariela Jaffé, Michael Krauthammer, Melinda Lohmann, Anna Mätzener, Sophie Mützel, Liliane Obrecht, Nicole Ritter, Matthias Spielkamp, Stephanie Volz

Dieses Positionspapier wurde im Rahmen eines Workshops erarbeitet, der vom 26.–28. August 2021 in Balsthal durchgeführt und vom Strategy Lab der Digital Society Initiative (DSI) der Universität Zürich finanziert wurde. Neben den Autor*innen dieses Papiers haben auch drei Vertreter*innen der Bundesverwaltung an diesem Workshop teilgenommen, nämlich Monique Cossali Sauvain (BJ), Roger Dubach (EDA) und Thomas Schneider (BAKOM). Sie vertreten die Schweiz im Ad Hoc Komitee des Europarates zu Künstlicher Intelligenz (CAHAI). Weitere Informationen: dsi.uzh.ch/strategy-lab

Die grossen technischen Fortschritte im Bereich der **Künstlichen Intelligenz (KI)** und der Einsatz dieser Technologien in einer Vielzahl von Bereichen werfen grundlegende Fragen zu den Auswirkungen auf Individuen und die Gesellschaft auf. Der Begriff der Künstlichen Intelligenz weckt bisweilen irreführende Assoziationen und diffuse Ängste. Aus technischer Perspektive handelt es sich um einen etablierten Sammelbegriff, der **eine Reihe von Technologien** umfasst, die automatisierte Entscheidungen fällen, Empfehlungen machen, Schlussfolgerungen ziehen oder Vorhersagen treffen. Dazu gehören wissensbasierte Systeme und statistische Methoden ebenso wie Ansätze des maschinellen Lernens (z.B. unter Einsatz neuronaler Netze). Die grosse Leistungsfähigkeit dieser Technologien basiert meist auf der Aneinanderreihung einer Vielzahl von mathematischen Optimierungen, die unter Nutzung grosser Rechnerkapazitäten Strukturen aus grossen Datenmengen extrahieren.

Um irreführende Assoziationen zu vermeiden, verwenden wir in diesem Positionspapier nicht den Begriff der Künstlichen Intelligenz (KI), sondern sprechen von «**algorithmischen Systemen**». Damit werden nicht bestimmte heutige oder künftige Technologien bezeichnet, sondern es wird auf die **Anwendung dieser Technologien in einem sozialen Kontext** verwiesen. Denn Bedarf nach einer rechtlichen Erfassung entsteht erst, wenn Technologien eingesetzt werden und Wirkung für Individuen und/oder die Gesellschaft entfalten. Der Begriff der algorithmischen Systeme erlaubt zudem, auch Anwendungen zu erfassen, die gleiche Wirkungen entfalten wie Künstliche Intelligenz, aber auf anderen Technologien beruhen.

Bei der Frage nach dem Regelungsbedarf ist zu beachten, dass der Einsatz von algorithmischen Systemen in der Regel **nicht zu völlig neuen Herausforderungen führt**. Einige davon bestehen auch, wenn keine algorithmischen Systeme verwendet werden, sondern Entscheide von Menschen getroffen werden – sie werden durch die Nutzung dieser Systeme nur besser sichtbar. Andere Herausforderungen wiederum erhalten durch die Nutzung solcher Systeme eine neue Qualität und Dimension, weil bspw. bestimmte Formen der Verhaltensbeeinflussung viel effizienter genutzt werden können – sowohl bezüglich der Präzision (z.B. zur Personalisierung) als auch hinsichtlich der Quantität (Skalierung).

Die **Europäische Kommission** hat am 21. April 2021 einen Vorschlag für eine Verordnung über Künst-

liche Intelligenz («AI Act») veröffentlicht[1], der nun Parlament und Ministerrat vorgelegt wird. Der **Europarat** hat eine erste Empfehlung zu Künstlicher Intelligenz verabschiedet[2] und einen Expert*innenausschuss (Ad hoc Committee on Artificial Intelligence, CAHAI) eingesetzt, der die Machbarkeit und mögliche Elemente eines Rechtsrahmens für die Entwicklung, Gestaltung und Anwendung von KI untersucht. Die Schweiz ist nicht an die Vorgaben der EU gebunden und es ist derzeit noch offen, ob sie eine allfällige Konvention des Europarates unterzeichnen wird. Absehbar ist immerhin, dass allfällige Vorgaben des Europarates den Mitgliedstaaten viel Freiraum bei der Ausgestaltung ihrer nationalen Lösungen lassen werden. **Die Schweiz sollte diesen Freiraum nutzen, um einen eigenen Ansatz zu entwickeln**. Dabei wird im Einzelnen zu entscheiden sein, welche Ansätze des EU-Rechts übernommen werden und wo die Schweiz zum Nutzen von betroffenen Personen, Wirtschaft und Gesellschaft vom EU-Recht bewusst abweichen sollte.

Dieses Positionspapier legt dar, welche **Ansätze zur rechtlichen Erfassung algorithmischer Systeme** in der Schweiz verfolgt werden sollten, welche Fragen besondere Beachtung erfordern und wie sich die Schweiz im Umfeld der europäischen Regulierungstendenzen positionieren soll.

Die Diskussion hat eine praktische und strategische Dringlichkeit, weil algorithmische Systeme zunehmend Einfluss auf das private und öffentliche Leben haben, in der Schweiz und im Ausland vermehrt Infrastrukturen für algorithmische Systeme geschaffen werden und sich das europäische und internationale Umfeld zunehmend der Regulierung dieser Systeme zuwendet, was unvermeidlich einen Einfluss auf die Schweiz nach sich ziehen wird.

1 Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Gesetz über künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union, COM(2021) 206 final.

2 Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems (Adopted by the Committee of Ministers on 8 April 2020 at the 1373rd meeting of the Ministers' Deputies) https://search.coe.int/cm/pages/result_details.aspx?ObjectId=09000016809e1154

Regelungsziele

Die rechtliche Erfassung der Herausforderungen des Einsatzes von algorithmischen Systemen dient zwei gleichwertigen Zielen: Zum einen soll die Regelung möglichst viel **Raum für die Entwicklung und Verwendung von algorithmischen Systemen** belassen, die für Einzelne und die Gesellschaft einen Nutzen schaffen. Zum andern ist sicherzustellen, dass die von der Verwendung von algorithmischen Systemen betroffenen Personen und die Gesellschaft als Ganzes aus diesen Verwendungen **keine Nachteile** erleiden, dass also bspw. betroffene Personen nicht diskriminiert, Volksabstimmungen nicht manipuliert und rechtsstaatliche Prinzipien nicht untergraben werden.

Regelungsansatz

Der Einsatz von algorithmischen Systemen führt zu vielfältigen **Herausforderungen**, die mit den Mitteln des Rechts zu erfassen sind; im Vordergrund stehen **fünf Bereiche**: Erkennbarkeit und Nachvollziehbarkeit, Diskriminierung, Manipulation, Haftung sowie Datenschutz und Datensicherheit.

Die Herausforderungen, die durch algorithmische Systeme entstehen, sind vielfältig und weisen oft eine neue Dimension oder Qualität auf, sie bestehen aber nicht nur beim Einsatz solcher Systeme. Diese Herausforderungen sollten deshalb nicht durch ein generelles «KI-Gesetz» oder ein «Algorithmen-Gesetz» erfasst werden. Angezeigt ist vielmehr eine **Kombination von allgemeinen und sektorspezifischen Normen**. Dabei steht die **punktuelle Anpassung bestehender Gesetze** im Vordergrund. Denn die Rechtsordnung enthält bereits Normen, die in der Lage sind, viele der Herausforderungen zu erfassen, die mit dem Einsatz von algorithmischen Systemen verbunden sind. Allerdings dürfte es in nicht wenigen Fällen erforderlich sein, **die Auslegung und Anwendung bestehender Normen anzupassen**, um den neuen Herausforderungen angemessen begegnen zu können.

Angesichts der Vielzahl der Erscheinungsformen algorithmischer Systeme ist grundsätzlich ein **technologieneutraler Ansatz** zu wählen, der die Herausforderungen unabhängig von der Verwendung einer bestimmten Technologie zu erfassen vermag. Wegen der

raschen technischen Entwicklung kann eine Regelung nur Bestand haben, wenn sie nicht auf eine bestimmte Technologie ausgerichtet ist. Dieser Grundsatz gilt ohne Einschränkung für die Ausgestaltung allgemeiner Normen. Er schliesst aber die Fokussierung der Regulierung auf eine konkrete Technologie in spezifischen Sektoren (z.B. Medizinprodukte, Fahrzeuge) nicht aus.

Regelungsbedarf

Der Einsatz von algorithmischen Systemen ist in der Regel mit der Verarbeitung von Daten verbunden. Handelt es sich dabei um Personendaten, greift das **Datenschutzrecht**. Die Bearbeitung von Personendaten durch algorithmische Systeme wirft allerdings keine grundlegend neuen Fragen auf. Es erscheint deshalb grundsätzlich möglich, die Herausforderungen für den Schutz der Privatsphäre und den Datenschutz mit den Mitteln des bestehenden Datenschutzrechts zu lösen.

Die Nutzung algorithmischer Systeme führt allerdings auch zu weiteren Herausforderungen. So ist der Einsatz solcher Systeme für die Betroffenen oft nicht **erkennbar** und ihre Funktionsweise nicht **nachvollziehbar**. Zudem können solche Systeme Menschen **diskriminieren** und in ihrem Denken und Handeln **manipulieren**. Ausserdem wirft der Einsatz algorithmischer Systeme neue **Haftungsfragen** auf. In allen diesen Bereichen besteht Regelungsbedarf. Das gilt auch für die **Gewährleistung der Sicherheit autonomer Systeme** und für bestimmte **Zulassungsverfahren**. Schliesslich fragt sich, ob der Einsatz von bestimmten, besonders problematischen autonomen Systemen (zumindest einstweilen) verboten werden sollte.

Erkennbarkeit und Nachvollziehbarkeit

Der Einsatz von algorithmischen Systemen und deren Funktionsweise muss für die betroffenen Personen erkennbar und verständlich sein. Diese Transparenz hat mehrere Dimensionen:

(1) Personen, die mit algorithmischen Systemen interagieren, müssen erkennen können, dass sie dies mit einem solchen System und nicht mit einem Menschen tun. Dies kann durch Einführung einer **Kennzeichnungspflicht beim Einsatz von algorithmischen Systemen** erreicht werden. Da die Interaktion eines algorithmischen Systems mit einer Person in aller Regel mit der Bearbeitung von Personendaten verbunden ist, könnte eine solche Kennzeichnungspflicht im Datenschutzgesetz vorgesehen werden.

(2) Personen, die von der Entscheidung eines algorithmischen Systems in relevanter Weise betroffen sind, müssen diese **Entscheidung nachvollziehen** können. Das bedeutet nicht, dass die Personen die technische Funktionsweise der Systeme im Einzelnen verstehen müssen; die Nachvollziehbarkeit muss vielmehr adressatengerecht sein. Der Umfang der Nachvollziehbarkeit hängt zudem von der Bedeutung der Entscheidung für die betroffene Person und den rechtlichen Anforderungen (z.B. Begründung von Urteilen von Gerichten oder Verfügungen von Behörden) im konkreten Kontext ab. Sicherzustellen ist deshalb, dass die betroffenen Personen in der Lage sind, die einer automatisierten Entscheidung zugrundeliegende Logik (insb. verwendete Daten und für die Entscheidung relevante Kriterien) zu verstehen und die notwendigen Informationen zu erhalten, um die Entscheidung gegebenenfalls anzufechten. Diese Informationen müssen leicht zugänglich und für Laien verständlich verfügbar gemacht werden.

(3) Zusätzlich zur individuellen **Erkennbarkeit** ist beim staatlichen Einsatz algorithmischer Systeme die Erkennbarkeit für die **interessierte Öffentlichkeit** sicherzustellen. Denkbar wäre, hierzu ein öffentlich zugängliches Register zu schaffen, aus dem ersichtlich wird, in welchen Bereichen die öffentliche Verwaltung algorithmische Systeme einsetzt. Ein solches Register sollte unter anderem Auskunft geben über die Art und Herkunft der bearbeiteten Daten, die Rechtsgrundlage, den Zweck und die Mittel der Bearbeitung, das verantwortliche Organ, die Logik des algorithmischen Systems und die Akteure, die an der Entwicklung des Systems mitgewirkt haben. Diese Informationen sollten leicht zugänglich sein und in einem standardisierten Format aufbereitet werden.

Diskriminierung

Die Aufgabe von algorithmischen Systemen besteht oft darin, Unterscheidungen zu treffen. Diese Unterscheidungen sind dann problematisch, wenn **Personen aufgrund von geschützten Merkmalen** wie Herkunft, Rasse, Geschlecht, Alter, Sprache, soziale Stellung, Lebensform, religiöse, weltanschauliche oder politische Überzeugung oder körperliche, geistige oder psychische Behinderung **unterschiedlich behandelt** werden, ohne dass dafür ein sachlicher Grund besteht. In diesem Fall liegt eine Diskriminierung vor. Bei algorithmischen Systemen können Diskriminierungen namentlich vorkommen, weil sie direkt oder indirekt geschützte Merkmale als Entscheidungsparameter verwenden oder weil sie mit Daten trainiert werden, die einen «bias» aufweisen. So können bestimmte gesellschaftlich existierende Vorurteile in Prognosen oder Entscheidungen solcher Systeme reproduziert werden. In vielen Fällen macht der Einsatz von algorithmischen Systemen die Diskriminierung aber erst sichtbar. Damit eröffnet der Einsatz solcher Systeme auch die Möglichkeit, gegen Diskriminierungen vorzugehen.

Die Problematik der Diskriminierung geht weit über den Einsatz algorithmischer Systeme hinaus, wird durch deren Einsatz aber besonders deutlich. Das Problem der Diskriminierung sollte deshalb durch Regeln erfasst werden, die **unabhängig** davon greifen, ob die diskriminierende Entscheidung oder Handlung von einem Menschen oder einer Maschine vorgenommen wird. Die aktuelle Rechtslage in der Schweiz verbietet in den meisten Fällen nur die Diskriminierung durch staatliche Akteure. Doch viele algorithmische Systeme werden von Privaten eingesetzt, etwa bei der Kreditvergabe oder bei der Selektion von Bewerbungen. Diese Diskriminierungen könnten durch ein **allgemeines Gleichbehandlungsgesetz** verhindert werden, das Diskriminierungen durch Private, insb. Unternehmen, aufgrund bestimmter, geschützter Merkmale erfasst und sanktioniert.

Der Nachweis einer Diskriminierung ist oft schwer zu erbringen. Dieses Problem könnte durch eine **Beweislastumkehr** gelöst werden. Die angeblich diskriminierte Person müsste das Vorliegen einer Diskriminierung nur hinreichend glaubhaft machen und das Unternehmen müsste dann den Nachweis erbringen, dass die Entscheidung nicht auf einem geschützten Merkmal beruht. Der Einsatz algorithmischer Systeme kann sich dabei auch als vorteilhaft erweisen, weil es – anders als bei menschlichen Entscheidungen – grundsätzlich möglich ist, die für die Entscheidung genutzten Kriterien zu erkennen und den Nachweis zu erbringen, dass eine Entscheidung nicht auf geschützte Merkmale abstellt.

Manipulation

Algorithmische Systeme können das Denken und Handeln von Personen beeinflussen, die mit solchen Systemen interagieren. Typische Beispiele sind das Anzeigen bestimmter und das Unterdrücken anderer relevanter Inhalte auf Social Media und die Personalisierung von Angeboten oder Preisen. Die zielgerichtete Beeinflussung des Denkens und Handelns einer Person durch einen Dritten (Manipulation) ist allerdings ein weit verbreitetes Phänomen, etwa bei Werbung. Die Beeinflussung durch Dritte ist zwar stets ein **Eingriff in die Autonomie** der betroffenen Person. Art und Ausmass der Beeinflussung sind aber höchst unterschiedlich und in vielen Fällen ist eine Beeinflussung unproblematisch. Das gilt beispielsweise dann, wenn eine Beeinflussung unspezifisch und für die betroffene Person erkennbar ist, wie etwa bei den traditionellen Formen der politischen und kommerziellen Werbung.

Bei der rechtlichen Erfassung von problematischen Formen der Manipulation ist insbesondere zwischen Entscheidungen bzw. Handlungen von Individuen in ihren Rollen als Konsument*innen und als Staatsbürger*innen zu unterscheiden:

(1) **Bei der Manipulation von Staatsbürger*innen** im Kontext demokratischer Prozesse steht der Schutz der **demokratischen Willensbildung** im Vordergrund. Diese kann beim Einsatz von algorithmischen Systemen gefährdet sein, weil solche Systeme besonders effiziente und kaum erkennbare Formen der Verbreitung von einseitiger Information, Übertreibung und Lüge erlauben. Zudem ist es möglich, einzelnen Personen (oder kleinen Gruppen) individualisierte Inhalte anzuzeigen, um ihr Denken, ihre Meinungsbildung und ihr Stimmverhalten gezielt zu beeinflussen. Diese Individualisierung der In-

halte kann dazu führen, dass bestimmte Aussagen gar nicht zum Gegenstand der öffentlichen Debatte werden, in der sie in Frage gestellt und gegebenenfalls widerlegt werden können. Bei der demokratischen Willensbildung kommt der **Informations- und Meinungsfreiheit** zentrale Bedeutung zu. Diese sichert politischen Akteuren und der Bevölkerung einen grossen Freiraum beim Wahrnehmen und Verbreiten von Informationen, der für die öffentliche Meinungsbildung zentral ist und nur sehr zurückhaltend eingeschränkt werden darf. Entsprechend sollte die Regulierung algorithmischer Systeme vorab das Schaffen von Transparenz über Art und Ausmass der Verbreitung von allfällig fragwürdigen Inhalten zum Ziel haben (z.B. das Bekanntmachen der Kriterien, nach denen Facebook Inhalte anzeigt, unterdrückt oder als problematisch kenntlich macht), ohne die Aussagen selbst zu bewerten. Diese Bewertung muss dem ergebnisoffenen Prozess der öffentlichen Meinungsbildung überlassen bleiben. Nutzer*innen sollen zudem durch geeignete Massnahmen erkennen können, wie Inhalte durch algorithmische Systeme individualisiert werden, um eine Sensibilität dafür zu entwickeln, wie sie dadurch beeinflusst werden.

(2) Bei der **Manipulation von Konsument*innen** stehen der Schutz der **individuellen Entscheidungsfreiheit** und der Schutz des **funktionierenden Wettbewerbs** gleichgeordnet nebeneinander. Auch bei Konsument*innen kommt der Manipulation durch die Verbreitung von falschen oder irreführenden Informationen zentrale Bedeutung zu. Diese Art der Manipulation kann allerdings mit dem geltenden Wettbewerbsrecht (UWG) erfasst werden. Anderes gilt bei anderen Formen der Manipulation, bspw. beim laufenden Anzeigen neuer Inhalte auf Social-Media-Plattformen mit dem Ziel, die Konsument*innen möglichst lange auf der Plattform zu halten, um ihnen möglichst viel Werbung anzeigen zu können. Hier ist zu prüfen, ob Handlungsbedarf besteht. Dies könnte insbesondere bei vulnerablen Personen der Fall sein (z.B. bei suchtartigem Social Media Konsum von Minderjährigen).

Für beide Gruppen muss Manipulation nicht notwendigerweise als Vorgang rechtlich erfasst werden. Vielmehr kann es ausreichen, Möglichkeiten zu schaffen, die es erlauben, **Entscheidungen rückgängig zu machen**, wenn sie aufgrund einer Manipulation erfolgt sind. Für Konsument*innen wäre insb. die Einführung von Widerrufsrechten denkbar, wie sie schon heute bei Haustürgeschäften und Telefonverkäufen und – in der EU – auch allgemein beim sog. Fernabsatz (insb. E-Commerce) bestehen. Bei Abstimmungen besteht zudem schon heute die Möglichkeit der Anfechtung, wenn das Ergebnis bspw. durch die Verbreitung von falschen Informationen massgeblich beeinflusst wurde.

Haftung

Eine zentrale Herausforderung beim Einsatz algorithmischer Systeme ist die Haftung im Fall eines Schadens. Zwar finden die Normen des allgemeinen Haftpflichtrechts auch auf solche Systeme Anwendung; der Nachweis der Voraussetzungen für die **Haftung von Betreiber*innen** ist allerdings mit Schwierigkeiten verbunden, insbesondere beim Verschulden. In bestimmten Sektoren stehen bereits verschuldensunabhängige Haftungsregeln zur Verfügung, die auch bei algorithmischen Systemen greifen (z.B. für Fahrzeuge im Strassenverkehrsgesetz oder für Drohnen im Luftverkehrsgesetz). Von der Einführung einer allgemeinen Betreiberhaftung in Form einer Gefährdungshaftung ist zwar abzusehen. Zu prüfen ist aber, ob in **weiteren Sektoren für Betreiber*innen algorithmischer Systeme eine verschuldensunabhängige Betreiberhaftung** eingeführt werden sollte. Ein sektorspezifisches Vorgehen würde eine behutsame Abstimmung mit ex ante zu erfüllenden Sicherheitsvorschriften ermöglichen.

In den Vordergrund rücken wird sodann die **Haftung der Hersteller*innen**. Als problematisch erweist sich, dass das Produktehaftungsgesetz auf herkömmliche Produkte und damit grundsätzlich auf physische Gegenstände zugeschnitten ist, die nach ihrer Herstellung in Verkehr gebracht werden und von den Hersteller*innen nicht mehr beeinflusst werden können. Die Erfassung von algorithmischen Systemen durch das **Produktehaftungsgesetz** setzt voraus, dass solche Systeme überhaupt als Produkte anerkannt werden.

Sodann sollten die Hersteller*innen für sichere (Weiter-)Entwicklungen ihrer Produkte haften. Gleichzeitig müssen sie sich jedoch bei unsachgemässer Einflussnahme anderer Beteiligter entlasten können. Das Schweizer Produktehaftungsgesetz ist entsprechend zu aktualisieren.

Sicherheit

Algorithmische Systeme müssen **gängigen Sicherheitsstandards** genügen. Sie müssen also ausreichend robust sowie vor schädlichen Umwelteinflüssen und Bedienungsfehlern geschützt sein. Zudem muss ein ausreichender Schutz gegen Angriffe gewährleistet sein, wobei auch neuere Formen von Angriffen (z.B. Manipulation von Trainingsdaten) zu beachten sind. Die Strenge der Anforderungen hängt von den Anwendungsbereichen ab; so müssen etwa algorithmische Systeme, die Prozesse in kritischen Infrastrukturen kontrollieren (z.B. Stromversorgung), strengeren Kriterien genügen als solche, die beispielsweise einen Staubsaugerroboter steuern.

Soweit algorithmische Systeme Personendaten bearbeiten, sind die Bestimmungen des Datenschutzrechts anwendbar, die eine angemessene Datensicherheit verlangen. Diese Bestimmungen zielen allerdings in erster Linie auf den Schutz von Personendaten und erfassen die Systeme nur indirekt. Zudem finden sie keine Anwendung, wenn algorithmische Systeme keine Personendaten bearbeiten, was gerade bei kritischen Infrastrukturen der Fall sein kann. Es ist deshalb zu prüfen, ob die Einführung eines allgemeinen **IT-Sicherheitsgesetzes** erforderlich ist. Als Alternative zu einer staatlichen Regulierung konkreter Sicherheitsanforderungen könnte sich die Allgemeinverbindlicherklärung von Standards aufdrängen, die von Standardisierungsorganisationen entwickelt werden.

Zulassungsverfahren

Bereits heute gibt es Produkte, die nur nach Zulassung durch eine staatliche Behörde auf den Markt gebracht werden dürfen (z.B. Fahrzeuge oder Medizinprodukte). Diese Zulassungsverfahren müssen auch dann durchlaufen werden, wenn Produkte algorithmische Systeme verwenden.

Bei den **bestehenden Zulassungsverfahren** sind die relevanten Voraussetzungen und Verfahren so anzupassen, dass sie die erforderliche Sicherheit und Qualität der Produkte auch dann gewährleisten, wenn diese auf dem Einsatz algorithmischer Systeme beruhen. Dabei ist zu beachten, dass algorithmische Systeme nach der Zulassung weiterentwickelt werden können oder sich gar selbst weiterentwickeln können (durch maschinelles Lernen). In diesen Fällen muss sichergestellt werden, dass die Zulassung bei jedem relevanten Entwicklungsschritt erneut überprüft wird («*life cycle regulation*»).

Zudem ist zu prüfen, ob **neue Zulassungsverfahren** geschaffen werden müssen, um die Sicherheit von risikobehafteten Produkten oder Diensten zu gewährleisten, die algorithmische Systeme verwenden. Im Vordergrund stehen dabei Systeme, die mit ihrer Umwelt interagieren, bspw. Pflege- oder Putzroboter, aber auch Spielzeuge. Zum anderen könnten auch Prognoseinstrumente, die in sensiblen Bereichen, etwa in der Strafverfolgung oder bei der Kriminalprävention, eingesetzt werden, einer Zulassung unterstellt werden. Bei weniger risikobehafteten Produkten könnte auch eine Zertifizierung vorgesehen werden.

Verbotene Anwendungen

Schliesslich ist zu prüfen, ob bestimmte Anwendungen von algorithmischen Systemen zu verbieten sind, weil sie zu Eingriffen in Grundrechte führen (oder führen können), die nicht hingenommen werden sollten. Als Alternative zu einem **Verbot** könnte auch ein **Moratorium** für den Einsatz bestimmter algorithmischer Systeme erlassen werden. Ein solches Moratorium würde es ermöglichen, die mittel- und langfristigen Folgen des Einsatzes von algorithmischen Systemen in kritischen Bereichen näher zu untersuchen und erst später zu entscheiden, ob der Einsatz solcher Systeme zugelassen werden soll. Im Vordergrund stehen aus heutiger Sicht die folgenden Anwendungen:

– Der Einsatz von **Gesichtserkennung und anderen biometrischen Fernerkennungsverfahren** im öffentlichen Raum, sofern die Gefahr besteht, dass diese algorithmischen Systeme für eine Massenüberwachung eingesetzt werden;

– Der Einsatz von **Social Scoring** mit dem Ziel, den Zugang zu grundlegenden Ressourcen (staatliche Dienstleistungen, Kredite, soziale Sicherheit etc.) zu regulieren.

Mit Blick auf die rasche technische Entwicklung ist zudem regelmässig zu evaluieren, ob neue Formen der Nutzung von algorithmischen Systemen (z.B. zur autonomen Ausübung tödlicher Gewalt im Sicherheitsbereich) ebenfalls verboten werden sollten.

Position der Schweiz im internationalen Umfeld

Aktuell wird in verschiedenen Rechtsräumen (EU, USA, China) an der Regulierung von algorithmischen Systemen gearbeitet. Relevant für die Schweiz sind insbesondere die Entwicklungen in der EU und im Europarat. Die Schweiz sollte **keine passive Übernahme dieser Regulierungsansätze** anstreben. Vielmehr sollte sie – basierend auf den in diesem Positionspapier formulierten Grundsätzen – eine eigene Position erarbeiten und diese aktiv zusammen mit internationalen Partner*innen mit ähnlichen Vorstellungen in den internationalen und insbesondere europäischen Diskurs einbringen. Dabei sollte die Kohärenz von Innen- und Aussenpolitik gewahrt bleiben und der aktive Diskurs auch innenpolitisch gespiegelt werden.

Schweizer Unternehmen, die autonome Systeme auf dem **europäischen Markt** anbieten oder einsetzen wollen, werden sich an die künftigen Vorgaben des EU-Rechts halten müssen. Das bedeutet aber nicht, dass die Schweiz diese Vorgaben in ihr nationales Recht übernehmen sollte. Vielmehr scheint es sinnvoll, durch einen hinreichend offenen rechtlichen Rahmen (bspw. durch ein allgemeines Verbot von Diskriminierung statt durch spezifische Vorgaben zu Risikomanagement und Datenqualität) Spielraum für diejenigen Schweizer Unternehmen zu schaffen, die ihre Produkte (noch) nicht auf dem europäischen Markt anbieten wollen.

Weiteres Vorgehen

Dieses Positionspapier zeigt, dass in der Schweiz Handlungsbedarf besteht. Die mit dem Einsatz von algorithmischen Systemen durch Unternehmen und den Staat verbundenen Herausforderungen sind hinreichend deutlich erkennbar. Vor diesem Hintergrund und mit Blick auf die Entwicklungen im Ausland sollte die Schweiz **zeitnah mit der Erarbeitung von Normen** beginnen, welche die skizzierten Herausforderungen angemessen erfassen können. Diese Arbeit sollte von einer breit aufgestellten, interdisziplinär zusammengesetzten **Expert*innenkommission** übernommen werden. In vielen Bereichen besteht zudem noch **Forschungsbedarf**, bspw. im Bereich der Manipulation. Die erforderlichen Forschungsarbeiten sollten mit hoher Intensität parallel zur Arbeit einer Expert*innenkommission fortgeführt werden, um sicherzustellen, dass die Regelung der Schweiz auf gesicherten wissenschaftlichen Grundlagen aufbauen kann.