

Fabienne Graf / Liliane Obrecht

## **Rechtliche Rahmenbedingungen für Künstliche Intelligenz in der Schweiz**

### **Tagungsbericht der Veranstaltung «Rechtliche Rahmenbedingungen für Künstliche Intelligenz in der Schweiz»**

---

Künstliche Intelligenz (KI) gilt als Schlüsseltechnologie zur Lösung gesellschaftlicher Herausforderungen. Bevor sie jedoch nachhaltig eingesetzt werden kann, muss sie in einen rechtlichen Kontext gestellt werden. Erste Ansätze wurden mit der Veranstaltung «Nachvollziehbare Algorithmen: ein Rechtsrahmen für den Einsatz Künstlicher Intelligenz» vorgestellt. Die Referierenden diskutierten die mit KI verbundenen Herausforderungen und Rechtsfragen. Der Fokus lag auf der Schweiz sowie ihrer Positionierung im europäischen und internationalen Regelungsumfeld. Auf die Referate folgte eine rege Podiumsdiskussion, wobei diverse Stimmen der unterschiedlichsten Interessengruppen Gehör fanden und Fragen aus dem Publikum aufgegriffen wurden.

---

Beitragsart: Tagungsberichte

Rechtsgebiete: Informatik und Recht, Privatrecht, Verwaltungsrecht, Europäisches Wirtschaftsrecht

Zitiervorschlag: Fabienne Graf / Liliane Obrecht, Rechtliche Rahmenbedingungen für Künstliche Intelligenz in der Schweiz, in: Jusletter 29. November 2021

## Inhaltsübersicht

- I. Rahmen der Veranstaltung
- II. Herausforderungen beim Einsatz von Künstlicher Intelligenz
- III. Internationale Entwicklungen zur Regulierung von KI
  - 1. Bisheriger internationaler Diskurs
  - 2. KI im internationalen Diskurs
- IV. KI-Verordnungsentwurf: Perspektive der Informatik
- V. Positionspapier Künstliche Intelligenz
- VI. Grundzüge einer möglichen Regulierung
  - 1. Manipulation
  - 2. Nachvollziehbarkeit (Perspektive: öffentliche Verwaltung)
- VII. Panel- und Plenumsdiskussion
- VIII. Würdigung und Ausblick

### I. Rahmen der Veranstaltung

[1] Am 10. November 2021 fand an der Universität Zürich eine gemeinsame Veranstaltung des Projekts «Nachvollziehbare Algorithmen: ein Rechtsrahmen für den Einsatz von Künstlicher Intelligenz» und des Strategy Lab der Digital Society Initiative (DSI) statt. Das erwähnte, von der Mercator-Stiftung finanzierte Forschungsprojekt wird vom Center for Information Technology, Society, and Law (ITSL) der Universität Zürich in Zusammenarbeit mit der Universität Basel geführt. Die Leitung haben Prof. Dr. NADJA BRAUN BINDER (Universität Basel) und Prof. Dr. FLORENT THOUVENIN (Universität Zürich, ITSL) inne.

[2] Eine engagierte Auseinandersetzung mit dem Thema der Künstlichen Intelligenz (nachfolgend: KI; englisch: *Artificial Intelligence, AI*) ging der Veranstaltung bereits im Rahmen des DSI Strategy Lab 2021 voraus. Sie mündete in der Erarbeitung eines Positionspapiers im Rahmen eines zweitägigen interdisziplinären Workshops im August 2021. Inhaltlich bildete diese Stellungnahme neben dem lancierten Forschungsprojekt den zweiten Schwerpunkt der Veranstaltung. Strategisch soll das Positionspapier den Ausgangspunkt für eine produktive politische Debatte bilden und die Legislative motivieren, sich der Fragen im Bereich der KI verbindlich anzunehmen.

### II. Herausforderungen beim Einsatz von Künstlicher Intelligenz

[3] In der Begrüssung identifizierten NADJA BRAUN BINDER und FLORENT THOUVENIN gestützt auf den aktuellen Stand der Forschung fünf zentrale Bereiche, in denen ein Einsatz von KI mit besonderen Herausforderungen verbunden ist: Transparenz, Privatsphäre, Diskriminierung, Manipulation und Haftung.

[4] Daran anknüpfend erörterte PD Dr. MARKUS CHRISTEN, Geschäftsführer der Digital Society Initiative (DSI) der Universität Zürich, diese Herausforderungen. Dabei konnte die Frage «Was ist eigentlich KI?» nicht abschliessend beantwortet werden, denn KI könne viele verschiedene Technologien umfassen. Für viele basiere das heutige Verständnis von KI primär auf künstlichen neuronalen Netzwerken (englisch: *neural networks*). Dabei handle es sich um eine Basistechnologie, die in vielen unterschiedlichen Bereichen eingesetzt werden könne. Diese Bereiche strukturierten sich nach einem «Sense-Think-Act-Paradigma», wobei sich KI bspw. zur Objekterkennung («Sense»), in Form Maschinellen Lernens («Think») oder im Kontext von Deep Fakes («Act») ein-

setzen liesse. CHRISTEN erörterte in einem weiteren Schritt die aktuell häufig diskutierten Fragen und würdigte deren gesellschaftliche Perspektive ebenso wie individuelle Applikationen. Die Fragen umfassten die bereits erwähnte Transparenz, die Autonomiefähigkeit technischer Systeme, das Sicherstellen menschlicher Kontrolle sowie die Forderung nach einem robusten Design von KI. Diese sehr grundsätzlichen Herausforderungen stellten sich zuvor bereits im Kontext menschlicher Handlungen. Auch geostrategische Elemente seien nicht zu vernachlässigen. So sei im Bereich der KI ein Vorpreschen der Supermächte als Herausforderung zu betrachten, zumal diese bereits erhebliche finanzielle Mittel in die Entwicklung von KI-Technologien eingesetzt hätten. Als Antwort habe die *European Group on Ethics in Science and New Technologies (EGE)* Anforderungen an mögliche Einsatzkonstellationen erarbeitet. Dabei werde augenscheinlich, dass die Technologien ein sehr breites Anforderungsspektrum abdecken müssen, was laut CHRISTEN noch viele Unsicherheiten berge.

### III. Internationale Entwicklungen zur Regulierung von KI

#### 1. Bisheriger internationaler Diskurs

[5] Botschafter THOMAS SCHNEIDER, Vizedirektor und Leiter der internationalen Beziehungen des Bundesamts für Kommunikation (BAKOM), begann sein Referat mit dem Vorstellen des Ad-hoc-Expertenausschusses für künstliche Intelligenz (CAHAI) des Europarates, in dessen Leitungsgremium (Bureau) er Mitglied ist. Der 2019 ins Leben gerufene Ausschuss sollte die Möglichkeit der Erarbeitung eines rechtlichen Rahmens für die Nutzung von KI unter Berücksichtigung von Menschenrechten, Demokratie und Rechtsstaatlichkeit prüfen. Die Mitarbeit im CAHAI stehe nicht nur den 47 Mitgliedsländern des Europarats, sondern auch anderen Staaten und Nichtregierungsorganisationen offen. Zudem seien offene Konsultationen durchgeführt worden. Ende 2020 veröffentlichte der CAHAI eine *«Feasibility Study»*, in welcher vorgeschlagen worden sei, dass der rechtliche Rahmen eine bindende Rahmenkonvention mit Grundprinzipien sowie weitere bindende und nichtbindende sektorielle Normen beinhalten solle. In seiner Schlussitzung im Dezember 2021 werde der CAHAI ein Dokument mit möglichen Elementen eines bindenden Instrumentes und dessen Umsetzungsmechanismen sowie Vorschläge für spezielle Anforderungen für die Nutzung von KI durch die öffentliche Hand verabschieden. Auf dieser Basis sollen dann ab Mitte 2022 Verhandlungen über ein bindendes Instrument beginnen. Dabei sei die Kooperation mit weiteren internationalen Organisationen, der EU sowie nicht-europäischen Ländern unabdingbar. Neben dem CAHAI existierten weitere Ausschüsse des Europarates, die sich mit KI befassten und in ihren Zuständigkeitsbereichen bereits *Soft Law* produziert hätten. Als Beispiel nannte SCHNEIDER den Lenkungsausschuss für Medien und die Informationsgesellschaft (CDMSI), der sich mit den Auswirkungen der Digitalisierung und von neuen Technologien auf Medien, Öffentlichkeit und Menschenrechte beschäftigt und dessen Vizevorsitz er derzeit inne hat.

[6] Ergänzend verwies SCHNEIDER auf Ansätze weiterer internationaler Organisationen, in denen das BAKOM aktiv mitarbeite. So habe die OECD bereits im Jahr 2019 erste Prinzipien zu KI verabschiedet, welche inklusives Wachstum und nachhaltige Entwicklung, menschliche Werte und Fairness, Transparenz, Sicherheit und Verantwortlichkeit umfassen. Noch im November 2021 werde die UNESCO eine Empfehlung zur Ethik von KI durch die UNESCO und somit das weltweit erste globale *Soft Law*-Instrument zu KI verabschieden. Die ITU organisiere seit

2017 die KI-Austauschplattform «*AI for Good*» zur Förderung von Partnerschaften und konkreten aktionsorientierten Projekten. Dabei folge sie den nachhaltigen Entwicklungszielen bis zum Jahr 2030. SCHNEIDERS Schlussbemerkung galt den Arbeiten im Bund. Er hielt fest, dass zwar noch keine Rechtsnormen erlassen worden seien, dass sich diverse Bundesstellen in ihren Bereichen intensiv mit den Herausforderungen zu KI auf nationaler und internationaler Ebene beschäftigten. So habe zum Beispiel das BAKOM zuhanden des Bundesrates einen Bericht zu möglichen Gouvernanz-Ansätzen für KI-basierte Intermediäre und Kommunikationsplattformen erarbeitet, welcher noch im November 2021 verabschiedet und veröffentlicht werden solle.

## 2. KI im internationalen Diskurs

[7] Botschafter ROGER DUBACH, stellvertretender Direktor der Direktion für Völkerrecht des Eidgenössischen Departements für auswärtige Angelegenheiten (EDA), präsentierte die Gründe, weshalb KI überhaupt ein internationales Thema darstellt. Die internationale Diskussion könne dabei strategische Motive beinhalten – auch in der Thematisierung grundlegender Anwendungen zur Ermöglichung von KI. Mit dem Technologieexport seien zudem Wertekollisionen aufgrund unterschiedlicher Moral- und Sittlichkeitsvorstellungen möglich. Es erneuere sich auch die Grundsatzfrage nach dem Verhältnis zwischen Mensch und Maschine.

[8] Aus den internationalen Diskussionen identifizierte DUBACH den Konsens über fünf Prinzipien: Transparenz, Fairness (oder Nicht-Diskriminierung), Nicht-Schaden (*non-maleficence*), Verantwortlichkeit und den Schutz der Privatsphäre. Sie seien den in der Schweiz vorherrschenden Vorstellungen sehr ähnlich. Ein zentraler Unterschied bestehe hingegen in der Präferenzordnung um eine horizontale oder sektorspezifische Regelung sowie der daraus abgeleiteten Herausforderung der Technologieneutralität. Während im internationalen Kontext allgemeine Forderungen nach horizontalen Regulierungen bestünden und neue Werkzeuge für KI erstellt werden sollten, sei die Stossrichtung in der Schweiz eine andere. Hier sollten bestehende Normen neu ausgelegt, einzelne Gesetze korrigiert und um spezifische Normen ergänzt werden. Die Technologieneutralität werde in der Schweiz in besonderem Masse gefordert. Anders die EU: sie entwickle von Grund auf neue Rechtsinstrumente für die «Technologie KI». Im internationalen Kontext sollten ausserdem private Personen an dieselben Pflichten gebunden werden wie der Staat. Diese Ansätze würden in der Schweiz so nicht verfolgt. Hier bestehe ein Spannungsverhältnis auf internationaler Ebene und ein Koordinationsbedarf. Der Bund befinde sich dazu in einem Findungsprozess. Doch eine gemeinsame Lösungsfindung sei für die Zukunft der Gesellschaft unerlässlich.

## IV. KI-Verordnungsentwurf: Perspektive der Informatik

[9] Das vierte Referat hielt Prof. Dr. BETTINA BERENDT, Professorin an der TU Berlin sowie der KU Leuven und Direktorin des Weizenbaum-Instituts für die vernetzte Gesellschaft. Sie war aus Berlin zugeschaltet und referierte zum KI-Verordnungsentwurf der EU-Kommission (KI-Verordnungsentwurf). Dieser wurde im April 2021 dem Europäischen Parlament vorgelegt.

[10] BERENDT begann mit der Unterscheidung zwischen einer spezifischen Regulierung von Anwendungen, wie sie im KI-Verordnungsentwurf der EU vorgeschlagen wird, und der Möglichkeit einer Regulierung über ein breiteres Systemdenken. Die Referentin positionierte sich nicht allein für einen einzigen Ansatz. Vielmehr betonte sie, dass die methodische Diskussion genutzt wer-

den solle, um über alte Probleme neu nachzudenken. Denn grundsätzlich solle eine Regulierung über Prinzipien stattfinden, die sich auf wissenschaftliche Erkenntnisse stützten und evidenzbasiert seien. Zudem erwähnte sie, dass aus Sicht der Informatik die Begrifflichkeit «Künstliche Intelligenz» wohl unzutreffend sei, passender hingegen «algorithmische Systeme».

[11] Ein Vergleich zwischen den Erwägungsgründen der DSGVO und des KI-Verordnungsentwurfes der EU verdeutliche, dass zwei Prinzipien der genaueren Betrachtung bedürften: die Transparenz und das Diskriminierungsverbot. Beide Erlasse gingen implizit von einer möglichen Diskriminierung aus. Deshalb bedürfe es Regeln zur Bearbeitung persönlicher Daten ebenso wie zu KI-Technologien. Daraufhin beleuchtete BERENDT die Massnahmen, die der KI-Verordnungsentwurf zur Verhinderung der Diskriminierung einsetzt. Mit dessen risikobasiertem Ansatz und der auf seiner Grundlage behaupteten Verhältnismässigkeit sei die Annahme verbunden, dass die Autonomie der (potenziellen) Anbieterinnen und Anbieter umso stärker eingeschränkt werden dürfe, je gefährlicher ein KI-System sei – und auch die Annahme, dass die geforderten Massnahmen geeignet seien, um die gefährdeten Grundrechte zu schützen. Das schlichte Verbot einer KI-Anwendung sei eine triviale Lösung, sofern überhaupt umsetzbar. Konkreter ging die Referentin auf die «erlaubten» Hochrisiko-KI-Systeme ein. Diese bedürften zu ihrem Einsatz Transparenz sowie des weiteren hoher Datenqualität, menschlicher Kontrolle, Sicherheit und Präzision. BERENDT hinterfragte daraufhin den Begriff der «Transparenz» und dessen Geeignetheit. Diese Kritik bildete den Ausgangspunkt für zwei Gedankenexperimente: Das erste stütze sich auf datenschutzrechtliche Erfahrungen wie die Einführung des informationellen Selbstbestimmungsrechts durch das deutsche Bundesverfassungsgericht. Das Wissen der Bürgerinnen und Bürger «wer was wann und bei welcher Gelegenheit über sie weiss», als breit angelegte Transparenz, sei zwar eine grundlegende Voraussetzung für das Funktionieren einer demokratischen Gesellschaft, jedoch sei es nicht die einzige. BERENDT betonte, dass das Grundrecht auf informationelle Selbstbestimmung neben der Transparenz insbesondere Kontrollbefugnisse und das grundsätzliche Bestimmen über die eigenen Daten vermittele. Dementsprechend enthalte das Datenschutzrecht viele weitere, datenverarbeitungsspezifische Prinzipien. Aus langjähriger evidenzbasierter Forschung ergebe sich, dass auch der KI-Bereich nicht ausschliesslich mittels Transparenz reguliert werden könne.

[12] Das zweite Gedankenexperiment illustrierte die Limitationen der Transparenz weiter. Auch wenn alles dokumentiert werde, alle Systeme hohe IT-Sicherheit aufwiesen und die Prozesse unter menschlicher Aufsicht stünden: Löse Transparenz das Problem der Gerechtigkeit oder der Diskriminierung? BERENDT verneinte diese Frage klar. Denn Diskriminierung könne auch transparent, nachvollziehbar und verstehbar erfolgen. Allerdings böten unterschiedliche Disziplinen Möglichkeiten für einen konstruktiven Umgang damit. Die Informatik sehe die Systementwicklung im Vordergrund.

## V. Positionspapier Künstliche Intelligenz

[13] Prof. ABRAHAM BERNSTEIN, Ph.D., Professor am Institut für Informatik und geschäftsführender Direktor der Digital Society Initiative (DSI) der Universität Zürich, stellte das Positionspapier des DSI Strategy Lab vor. BERNSTEIN hielt fest, dass es sich bei KI um eine Zukunftstechnologie handle, die ebenso die gesellschaftliche wie auch die persönliche Ebene erreiche. Dies veranschaulichten die bereits sehr aktiven internationalen Debatten. Die Regulationsspielräume

müssten vor den Umsetzungsschritten identifiziert werden. Mit einem solch umsichtigen Vorgehen werde das Potenzial von Technologien nicht verkannt und Innovationen weiter ermöglicht.

[14] Nach den einführenden Bemerkungen zur gesamtgesellschaftlichen Bedeutung ging BERNSTEIN auf den Regelungsansatz des Positionspapiers ein. Dabei seien drei Ansatzpunkte hervorzuheben: Erstens, das Abstellen auf den Begriff «algorithmische Systeme» (anstelle von «KI»). Zweitens solle die Technologieneutralität die Zukunftstauglichkeit der Regelung gewährleisten. Drittens rücke die sektorspezifische Anpassung bestehender Normen, anstelle einer horizontalen Regelung, in den Fokus.

[15] Der Regelungsbedarf wird im Positionspapier in fünf Kategorien unterteilt. Neben Erkennbarkeit und Nachvollziehbarkeit, Diskriminierung, Manipulation sowie die Haftung träten Datenschutz und Datensicherheit. Im Hinblick auf die Nachvollziehbarkeit forderte BERNSTEIN vor dem Hintergrund des Positionspapiers eine Kennzeichnungspflicht sowie die Erstellung eines öffentlichen Registers. So solle erkenntlich gemacht werden, wer die besagten Systeme wo einsetzt. Weiter müsse sichergestellt werden, dass keine Diskriminierung durch die algorithmischen Systeme stattfinden könne. BERNSTEIN betonte, dass diese Problematik unabhängig von der Entscheidungsform im privaten Bereich reguliert werden müsse, wobei ein allgemeines Gleichbehandlungsgesetz eine Möglichkeit darstelle. Da es für betroffene Personen schwierig sei darzulegen, dass tatsächlich eine Diskriminierung stattgefunden habe, könnte eine Beweislastumkehr in Erwägung gezogen werden. Als dritter Punkt wurde die Manipulation thematisiert. Diese habe zwei Dimensionen: Einerseits müsse die demokratische Willensbildung – und somit im weitesten Sinne die Meinungsfreiheit – geschützt werden. Andererseits stehe der Schutz der individuellen Entscheidungsfreiheit im Rahmen des marktwirtschaftlichen Wettbewerbs im Vordergrund. Dabei stehe das mögliche Rückgängigmachen im Fokus. Es handle sich dabei um Massnahmen, die es erlaubten, eine getroffene Entscheidung zurückzuziehen. Der vierte Punkt betraf die Haftungsfrage. Wer hafte, wenn ein Algorithmus falsche Ergebnisse wiedergebe? Abschliessende Gedanken zur Haftungsfrage mündeten in der Erkenntnis, dass algorithmische Systeme als Produkte erfasst werden müssten, um überhaupt eine Herstellerinnen- bzw. Herstellerhaftung zu ermöglichen. Neben genügender Sicherheitsstandards, die gegebenenfalls in der Erstellung eines IT-Sicherheitsgesetzes resultieren könnten, müsse die Erarbeitung neuer Zulassungsverfahren geprüft werden. BERNSTEIN präsentierte abschliessend das im Positionspapier vorgeschlagene Verbot bzw. Moratorium für gewisse Anwendungen. Darunter fielen erstens die Gesichtserkennung zum Zwecke der Massenüberwachung sowie zweitens Technologien, die ein *Social Scoring*-System zum Ziel hätten.

## VI. Grundzüge einer möglichen Regulierung

[16] Eine mögliche Regulierung umfasst sowohl privat- wie auch öffentlich-rechtliche Aspekte. Exemplarisch wurden von THOUVENIN und BRAUN BINDER die Manipulation und Nachvollziehbarkeit hervorgehoben.

### 1. Manipulation

[17] THOUVENIN wies einleitend auf die verbreitete Befürchtung hin, dass Menschen bei der Interaktion mit algorithmischen Systemen in einem diffusen Sinne in ihrem Denken und Handeln

manipuliert werden könnten. Dabei werde die Autonomie durch Dritte beeinträchtigt. Diese Beeinträchtigung könne zum Vorteil der betroffenen Person (*Nudge*) oder zu ihrem Nachteil sein. Damit stelle sich die zentrale Frage, ob die Beeinträchtigung der Autonomie als solche oder bloss die zum Nachteil der Betroffenen erfolgende Beeinträchtigung rechtlich zu erfassen sei. Das Phänomen der Manipulation sei jedenfalls vielgestaltig und schwer fassbar. Gerade im Bereich von KI sei es zentral, dass die effektiv bestehenden praktischen Probleme identifiziert werden, um eine passende rechtliche Erfassung von Manipulation zu ermöglichen.

[18] Zwei mögliche Konstellationen wurden von THOUVENIN unterschieden: die Manipulation von Staatsbürgerinnen und Staatsbürgern im Zusammenhang mit der politischen Willensbildung und die Manipulation von Konsumentinnen und Konsumenten im Zusammenhang mit Konsumentscheiden. Sein Referat fokussierte auf die zweite Konstellation. Nach verbreitetem Verständnis werde Manipulation anhand von vier Merkmalen definiert: Die absichtliche aber verdeckte Beeinflussung der Entscheidung einer anderen Person zum Vorteil des Manipulierenden, unter Ausnutzung einer gewissen Vulnerabilität (z. B. von Kindern oder Jugendlichen). Dieses Verständnis liege auch dem KI-Verordnungsentwurf der EU zugrunde, der in zwei Bestimmungen Fälle von Manipulation erfassen wolle. Bemerkenswert sei allerdings, dass der KI-Verordnungsentwurf nicht auf den Vorteil des Manipulierenden, sondern auf den Nachteil der Manipulierten abstelle. Aus rechtlicher Sicht erscheine dies durchaus sinnvoll, weil die Rechtsordnung in der Regel verhindern wolle, dass Dritten durch das Handeln eines Akteurs ein Nachteil entstehe, während das blosses Entstehen von Vorteilen beim handelnden Akteur in aller Regel als unproblematisch zu qualifizieren sei. Auch das Merkmal der Absicht erscheine aus (privat-)rechtlicher Sicht als problematisch, weil es in erster Linie auf die Wirkung einer Handlung ankomme und nicht auf die dahinterstehende Absicht; zudem seien subjektive Merkmale oft schwer festzustellen und zu beweisen. Der Fokus bei der rechtlichen Erfassung von Manipulation durch KI sollte deshalb auf dem Merkmal der verdeckten Beeinflussung der Entscheidung von Menschen und auf dem Nachteil liegen, der den Beeinflussten dadurch entstehen könne. Die Vulnerabilität könne kaum ein in allen Fällen notwendiges Anknüpfungsmerkmal sein, weil rechtlich in der Regel auf das Verständnis einer Durchschnittsperson abgestellt werde und auch nicht-vulnerable Personen manipuliert werden können. In gewissen Konstellationen könne die Vulnerabilität aber Anlass für den Erlass von Sonderregeln bilden, etwa beim Jugendschutz in den Sozialen Medien.

[19] THOUVENIN sprach sich dafür aus, die Manipulation von Konsumentinnen und Konsumenten grundsätzlich über das Bundesgesetz gegen den unlauteren Wettbewerb (UWG) zu erfassen. Dieses sei nicht nur hinreichend flexibel, um die vielfältigen Formen von Manipulation zu erfassen, sondern verfüge auch über eine reichhaltige Praxis, um Formen von Manipulation zu erfassen, die auf dem Verbreiten falscher oder irreführender Informationen beruhe. Das UWG sollte damit auch beim Einsatz algorithmischer Systeme ausreichenden Schutz bieten können. Schwieriger zu erfassen seien Konstellationen ausserhalb eines (direkten) Konsumkontextes, etwa das laufende Anzeigen von weiteren Inhalten auf Sozialen Medien mit dem Ziel, die Nutzerinnen und Nutzer möglichst lange auf einer Plattform zu halten. Zudem stelle sich stets die heikle Frage, wo die Grenze zwischen zulässiger Beeinflussung und unzulässiger Manipulation zu ziehen sei. Abstrakt lasse sich diese Grenze kaum definieren. Immerhin werde es aber Fälle geben, in welchen die Grenze der zulässigen Beeinflussung derart klar überschritten sei, dass der Bedarf nach einem Eingreifen des Rechts offensichtlich werde. Dabei müsse auch beachtet werden, dass Manipulation nicht nur in einem einzigen Moment auftrete, sondern Zeitdauer und Quantität einschlägige Kriterien für die Einordnung sein können.

## 2. Nachvollziehbarkeit (Perspektive: öffentliche Verwaltung)

[20] BRAUN BINDER eröffnete ihre Ausführungen mit der Feststellung, dass die öffentliche Verwaltung bei ihrer Tätigkeit Rechtsnormen anwende. Eine Rechtsnorm umfasse Tatbestand und Rechtsfolge. Der Tatbestand könne nach deskriptiven und normativen Tatbestandselementen unterschieden werden. Erstere verwendeten wirklichkeitsbezogene Begriffe (z. B. Mensch, Motorfahrzeug). Normative Elemente setzten bei nicht wirklichkeitsbezogenen Begriffen an, die auf gesellschaftliche Werte verwiesen (z. B. verhältnismässig, zumutbar). Sie seien weniger deutlich fassbar und würden von der Legislative eingesetzt, um behördlichen und richterlichen Spielraum zu ermöglichen. Gesetze blieben so über viele Jahre anwendbar. Ihre Bedeutung müsse durch Auslegung ermittelt werden. Bei der Rechtsfolge bestehe entweder kein Ermessensspielraum («wenn A, dann immer B») oder Rechtsfolgeermessen. Letzteres diene dazu, der Behörde Spielraum zu eröffnen, insbesondere, um dem konkreten Einzelfall Rechnung zu tragen. Allerdings seien dabei verschiedene Grundsätze wie das Rechtsgleichheitsgebot zu beachten.

[21] Algorithmische Systeme könnten in der öffentlichen Verwaltung nun bei genau dieser Normanwendung eingesetzt werden. Dabei entstünden allerdings juristische Herausforderungen, insbesondere bezüglich der Nachvollziehbarkeit einer behördlichen Entscheidung. Verschiedene Möglichkeiten seien denkbar: Ein algorithmischer Einsatz könnte bspw. nur dort erlaubt werden, wo Normen weder unbestimmte Rechtsbegriffe enthielten noch Rechtsfolgeermessen einräumten. Die Nachvollziehbarkeit könne in einer solchen Konstellation relativ einfach sichergestellt werden, da das Resultat der Rechtsanwendung mit dem gesetzlich vorgegebenen Resultat übereinstimmen müsse. Sofern algorithmische Systeme eingesetzt würden, um unbestimmte Rechtsbegriffe auszulegen bzw. Ermessensspielräume zu nutzen, sei die Nachvollziehbarkeit einer Entscheidung anderweitig sicherzustellen. Eine pauschale Erklärung, wie ein System funktioniert, genüge dabei nicht. Vielmehr sei die Entscheidung im konkreten Einzelfall nachvollziehbar zu machen. Ihr Referat endete mit einem Vorschlag zur Ergänzung des Art. 35 Abs. 1 Verwaltungsverfahrensgesetzes (VwVG) um einen Abs. 1<sup>bis</sup>: «Verfügungen, die mithilfe algorithmischer Systeme erlassen werden, sind speziell zu kennzeichnen und mit einer inhaltlich angemessenen Begründung zu versehen.» Die Kennzeichnungspflicht solle ein Mittel zur Transparenzherstellung sein, die Pflicht zur angemessenen Begründung den notwendigen Spielraum mit Blick auf unterschiedliche Anwendungskontexte belassen.

## VII. Panel- und Plenumsdiskussion

[22] Neben den Vortragenden kamen Prof. Dr. DIEGO KUONEN (Verantwortlicher *ad interim* des Kompetenznetzwerks Künstliche Intelligenz des Bundes CNAI, BFS), DANIEL SCHÖNBERGER (Head of Legal Schweiz, Google), Dr. des. ANGELA MÜLLER (Team Lead Policy & Advocacy, Algorithm-Watch Schweiz) sowie TORBEN STEPHAN (Leiter Digitalisierung und Gesellschaft, Stiftung Mercator Schweiz) zu Wort. Alle vier Personen hielten zunächst kurze Input-Referate, um die eigenen Perspektiven für eine gemeinsame Diskussion zu eröffnen.

[23] KUONEN betonte das Black Box-Problem von KI-Anwendungen. Er hielt dazu fest, dass das Problem nicht die Algorithmen, sondern die Daten seien. MÜLLER hob hervor, dass sich der Einsatz von algorithmenbasierten Systemen an gesellschaftlichen Grundwerten zu orientieren habe. Eine notwendige – wenn auch nicht hinreichende – Bedingung dafür sei Transparenz: Diese sei unabdingbar, um eine evidenzbasierte öffentliche Debatte und damit die Ausübung von in-

dividueller und demokratischer Kontrolle zu ermöglichen. SCHÖNBERGER gab Einblicke in den unternehmerischen Wandel bei Google von «*mobile first*» zu «*AI first*». Er betonte weiter, dass international inzwischen viele Rahmen zur Selbstregulierung bestünden. STEPHAN legte den Fokus wieder auf gesamtgesellschaftliche Werte. Er stellte in der Diskussion um KI ein noch immer starkes Schwanken «zwischen Weltuntergang und Allheilmittel» fest – wobei oft schlicht das Verständnis fehle.

[24] Die erste Frage der Paneldiskussion betraf die allgemeine Stossrichtung einer Regulierung von KI in der Schweiz: Muss die Schweiz auf rechtlicher Ebene etwas tun, um den Herausforderungen von KI zu begegnen – und wenn ja: was? SCHÖNBERGER verwies als erster auf den KI-Verordnungsentwurf der EU. Google beschäftige sich nicht primär mit der Frage nach der Gesetzgebung in der Schweiz, sondern fokussiere sich eher auf ihre Reaktion auf den KI-Verordnungsentwurf. SCHÖNBERGER sprach sich vor dem Hintergrund der ungelösten Haftungsfrage für eine sektorielle *ex ante* KI-Regulierung aus. Für MÜLLER ist es wichtig, dass bestehende Gesetze und Prinzipien auch auf automatisierte Entscheidungen angewendet werden. Rechtliche Lücken müssten zudem ausfindig gemacht und geschlossen werden. Einen Ansatz, der Systeme vorab in bestimmte Risikokategorien einteilt, lehne sie schon aus dem Grund ab, dass die Risiken wesentlich vom Kontext abhängen, in dem ein System eingesetzt wird. Die Vorstellung, dass die Risiken durch Massnahmen auf technologischer Ebene beseitigt werden könnten, übertrage zudem denjenigen Unternehmen viel Gestaltungsmacht, die die Systeme programmieren. Dies ergänzte THOUVENIN mit dem Vorschlag, problemkonzentriert und nicht technologiekonzentriert zu normieren. KUONEN erinnerte an die Datenqualität, die ausschlaggebend sei. STEPHAN fügte dem an, dass wohl der wichtigste Aspekt der Einbezug der Zivilgesellschaft sei. SCHNEIDER stellte grundsätzlich den langwierigen Prozess der Gesetzes- und Verfassungsrevisionen in Frage, der bei der rechtlichen Erfassung von neuen Technologien besonders problematisch sei.

[25] Inwiefern sind die Zivilgesellschaft und NGOs, wie die hier vertretene AlgorithmWatch Schweiz, relevant? MÜLLER verwies auf die bereits bestehende Zusammenarbeit von AlgorithmWatch auf EU-Ebene mit anderen Organisationen und den Austausch mit EU-Entscheidungsträgerinnen und Entscheidungsträgern. Gleichzeitig erinnerte sie an das direkt-demokratische System der Schweiz. Dieses vernetze verschiedene Interessengruppen und eröffne breite Partizipationsmöglichkeiten. Es müsse jedoch sichergestellt werden, dass auch die Perspektive der betroffenen Personen und Gruppen stets miteinfliesse. KUONEN ergänzte diese Ausführungen mit dem Hinweis auf das Kompetenznetzwerk KI des Bundes. Dabei handle es sich um eine Geschäftsstelle, die einerseits Expertinnen und Experten vereinen solle, um Wissen systematisiert akkumulieren und nutzen zu können. Andererseits solle sie ab Frühjahr 2022 auch bundesintern und -extern als Ansprechpartnerin für jegliche KI-Themen agieren.

[26] Die Diskussion wandte sich dann der Selbstregulierung zu. Zur Frage, ob diese nicht besser sei als ein KI-Gesetz, äusserte sich zuerst BRAUN BINDER. Eine juristische Positionierung bestehe bereits; ein KI-Gesetz werde abgelehnt – stattdessen seien punktuelle Regulierungen notwendig. BERNSTEIN ergänzte, dass beide Bereiche, die rechtliche Regulierung und die Selbstregulierung, ineinandergreifen müssten. Dazu hielt KUONEN fest, dass die Etablierung von Selbstregulierung wohl mindestens genauso lange dauern würde wie eine rechtliche Regelung. Auch STEPHAN kritisierte die Selbstregulierung, weil diese wohl kaum weltweit funktionieren würde. Als Beispiel zog er die unterschiedlichen Wertvorstellungen z. B. zwischen den USA und China heran. THOUVENIN führte das Gesagte zusammen: Es sei die Aufgabe der Rechtsordnung, Ziele vorzugeben, deren

Umsetzung durch konkrete Lösungen auf einer technischen Ebene dann im Rahmen einer Selbstregulierung erfolgen könne.

[27] Die Frage nach dem Bestehen eines Spannungsverhältnisses zwischen Wertvorstellungen war Thema des letzten Teils der Diskussionsrunde. Sind Bias und Diskriminierung, die von KI perpetuiert werden, nicht global verschieden? Auf die Frage, ob denn nicht auch Rechtsetzung und Rechtsprechung von Bias geprägt seien, knüpfte MÜLLER am Demokratieprinzip an; bei den zu erlassenden Normen werde es sich um im demokratischen Prozess hergeleitete Bestimmungen handeln. BERNSTEIN suchte nach einem Mittelweg, indem er diese Wertvorstellungen als mitunter heikles Feld der Diversität identifizierte. Auch er stellte auf das Demokratieverständnis ab und dieses Verständnis unterscheide sich geographisch. Schlussendlich handle es sich um «sehr subtile Dinge», welche die Gesellschaft mitgestalteten – und KI trete als solcher Aspekt hinzu.

### **VIII. Würdigung und Ausblick**

[28] Mit dem Abschluss der Veranstaltung ist die Diskussion um einen adäquaten Rechtsrahmen für KI in der Schweiz nun erst richtig eröffnet. Die an der Universität Zürich vorgestellten Positionen haben dazu einen umfassenden und wertvollen Überblick geschaffen. Ebenso deutlich wurde, dass im Hinblick auf den anstehenden Regelungsprozess noch zahlreiche offene Fragen bestehen. Diese zu adressieren und konstruktive Lösungsansätze zu erarbeiten, stellt eine breite und nicht bloss rechtliche Herausforderung dar. Ergänzend zum demokratischen Prozess und zivilgesellschaftlichen Engagement stellt sich die interdisziplinäre und rechtswissenschaftliche Forschung dieser Aufgabe. Auf der Suche nach einem ebenso nutzbringenden wie kritischen Umgang mit dem Phänomen der KI gibt es noch viel zu tun; diesem Handlungsauftrag sollte möglichst zügig gefolgt werden.

---

FABIENNE GRAF, MLaw, LL.M. (Duke), arbeitet als wissenschaftliche Mitarbeiterin am Center for Information Technology, Society, and Law (ITSL) der Universität Zürich sowie als wissenschaftliche Assistentin am Lehrstuhl für Wirtschaftsrecht und Rechtsphilosophie der Universität Luzern. Fabienne Graf ist Doktorandin der Humboldt-Universität zu Berlin und der Universität Luzern. Interessenbindung: Mitarbeit im Projekt «Nachvollziehbare Algorithmen: ein Rechtsrahmen für den Einsatz von Künstlicher Intelligenz», ITSL, Universität Zürich.

LILIANE OBRECHT, BLaw, studiert im Master Rechtswissenschaften an den Universitäten Basel und Lausanne und arbeitet als Studierende in Assistenzfunktion bei Prof. Dr. Nadja Braun Binder an der Universität Basel, wo sie demnächst ihr Doktoratsstudium beginnt. Interessenbindung: Mitarbeit im Projekt «Nachvollziehbare Algorithmen: ein Rechtsrahmen für Künstliche Intelligenz» sowie Mitautorin des Positionspapiers des DSI Strategy Lab.

Die Autorinnen danken allen Referierenden für die konstruktiven Ergänzungen und hilfreichen Anmerkungen im Text.

*Hinweis: Alle Tagungsbeiträge sind als Video-Podcast frei verfügbar unter: <https://www.itsl.uzh.ch/de/podcast-KI-2021> und [www.dsi.uzh.ch/strategy-lab](http://www.dsi.uzh.ch/strategy-lab).*