

Demokratie und Künstliche Intelligenz

1 Einleitung: Künstliche Intelligenz, ein neues Thema für die Politikwissenschaft

Welchen Beitrag kann die Politikwissenschaft zur öffentlichen Verständigung über das Phänomen der Künstlichen Intelligenz leisten? Angesichts des Aufstiegs statistischer Datenanalytik zu einer neuen Querschnittstechnologie im öffentlichen und privaten Leben widmen sich auch die sozialwissenschaftlichen Disziplinen in zunehmendem Umfang diesem Gegenstand. Der politikwissenschaftliche Fokus richtet sich hierbei auf die Beziehung zwischen den algorithmischen Analysen, die heute als „Künstliche Intelligenz“ bezeichnet werden und dem politischen System. Dazu gehören zum einen die Ideale, Normen und Prinzipien sowie die Institutionen und Prozesse der gelebten Demokratie, zum anderen aber auch die Arbeitsweisen der öffentlichen Verwaltung und des Rechtsstaats. Das Interesse gilt vor allem den Änderungen, die der Einsatz des maschinellen Lernens für das politische Gefüge mit sich bringt. In diesem Text wird es hauptsächlich um das Zusammenspiel von KI und demokratischer Selbstbestimmung, insbesondere um den Bereich der politischen Willensbildung und Partizipation gehen.

Obwohl KI noch ein recht randständiges Thema in der Politikwissenschaft ist, zeichnen sich bereits unterschiedliche Sichtweisen auf die Folgen des Einsatzes für die Demokratie ab. Darin drücken sich nicht zuletzt unterschiedliche Vorverständnisse algorithmischer Systeme und ihrer Wirkungs- bzw. Machtpotenziale aus. Schon deshalb scheint es sinnvoll, diesen Beitrag im nächsten Abschnitt mit einer kurzen Reflexion über die erstaunliche terminologische Unbestimmtheit von KI zu beginnen. Im Anschluss daran werden im dritten Abschnitt zwei politikwissenschaftliche Forschungsperspektiven vorgestellt: KI als Gegenspielerin zur oder aber als Mitspielerin in der Demokratie. Der zentrale Befund dieser Gegenüberstellung ist, dass sich algorithmische Systeme als gestaltbare Wissenstechnologien verstehen lassen, die ihr Verhältnis zur demokratischen Regierungsweise ausdrücklich nicht selbst vorgeben. Mit anderen Worten besteht keine Systemkonkurrenz zu den Institutionen der kollektiven Selbstbestimmung. Die spezifische Rolle und Wirkungsweise dessen, was heute als KI gilt, wird bislang vielmehr im Zuge ihres Einsatzes in einzelnen Anwendungskontexten entschieden.

Der erhebliche gesellschaftliche Einfluss auf die Funktionsweise von KI-Systemen lenkt den Blick auf die Rolle der politischen Partizipation, also der gesellschaftlichen Kräfte, die die Entwicklung dieser Technologien kritisch begleiten und zu prägen versuchen. Der vierte Abschnitt beschäftigt sich mit den spezifischen Stärken und Schwächen, die auch den schwierigen Rahmenbedingungen des politischen Engagements geschuldet sind. In der Praxis lässt sich der Einsatz von KI-basierten Verfahren nämlich kaum beobachten, weil sich weder staatliche noch private Anwender:innen gern in die Karten schauen lassen. Zu den Schwächen des politischen Engagements gehört aber auch der Fokus auf die Gegenwart und das Fehlen einer langfristigen Perspektive, die auch diejenigen Transformationsprozesse in den Blick nimmt, die die Veralltäglicung und Normalisierung dieser prädiktiven Technologien begleiten werden. Exemplarische Überlegungen dazu stellt der Exkurs an, der den Abschluss dieses Textes bildet.

2 Künstliche Intelligenz als flüchtiges Objekt

Die Verwendung des Begriffs der Künstlichen Intelligenz (KI) wird gegenwärtig vielfach unter Vorbehalt gestellt, weil es bisher nicht gelungen ist, diese Technologie präzise einzugrenzen oder gar zu definieren. Eine Ursache für die Unschärfe im Verständnis von KI liegt in den Irrungen und Wirrungen ihrer Entstehung. Was in den späten 1950er-Jahren als wissenschaftlicher Konflikt über unterschiedliche Programmierungsstrategien intelligenter Maschinen begann, ist unter dem Einfluss erheblicher wirtschaftlicher Investitionen in den 2010er-Jahren zugunsten des einstmals unterlegenen Ansatzes entschieden worden (vgl. Cardon & Cointet 2018). Als KI gilt heute, was noch in den frühen 2000er-Jahren als „maschinelles Lernen“, also das Erkennen von Mustern in großen Datensätzen, bezeichnet wurde (vgl. Jordan 2019). Der Begriff „KI“ war dagegen für ehrgeizigere Projekte zur Nachbildung der menschlichen Intelligenz reserviert. Das „rebranding“ von KI als maschinelles Lernen veranlasst manche Beobachter:innen, die Existenz oder Machbarkeit von KI grundsätzlich in Abrede zu stellen. Andere gehen pragmatischer vor und verwenden den Begriff für „whatever we are doing next in computing“ (Recker 2021: 5). Dritte wiederum verzichten auf die Verwendung des Begriffs und weichen stattdessen auf Abstraktionen wie „prädiktive Technologien“, „agentic machines“ oder „algorithmische Systeme“ aus (Joyce et al. 2021: 2; vgl. auch Dignum 2022). Der allgemeinen Attraktivität von KI können all diese Unklarheiten offenbar nichts anhaben, erlauben sie es doch, vollmundige Versprechen zu geben, ohne diese jemals einlösen zu müssen (Joyce et al. 2021: 2).

Einstweilen behauptet sich KI als „public facing term“ (ebenda), während die professionelle Fachgemeinde konkreteren Bezeichnungen wie „maschinelles Lernen“ oder „neuronale Netze“ den Vorzug gibt.

Der unbestimmte Realitätsbezug von KI erweist sich allerdings als Problem, wenn man ihre Wechselwirkungen mit Gesellschaft und Demokratie untersuchen möchte. Wie lässt sich das Zusammenspiel zwischen Demokratie und einer Technologie untersuchen, „über die keine Übereinstimmung besteht, weder unter Experten noch unter jenen, die von ihrem angeblich disruptiven Charakter betroffen sind“ (Hildebrandt 2020: 74, Original in Englisch)? Läuft man nicht Gefahr, Diagnosen zu stellen, die ungewollt zur Essenzialisierung und Normalisierung von bloßen Worthülsen beitragen? Die sozialwissenschaftliche Forschung zu KI reagiert auf dieses Problem mit unterschiedlichen Formen der Diskurskritik, deren Gemeinsamkeit vielleicht darin liegt, weniger Gewicht auf Definitionen zu legen und die Aufmerksamkeit stattdessen auf die Interpretationsweisen und ihre Wirkungen zu lenken.

Beispielhaft steht hierfür die Analyse von KI als soziotechnischem „Imaginary“ (Jasanoff & Kim 2015). So schlagen Natale und Ballatore (2020) vor, KI als machtvollen Mythos über die „denkende Maschine“ zu verstehen. Bei Mythen ginge es nicht in erster Linie um ihren Wahrheitsgehalt, sondern um ihre narrativen, „welterzeugenden“ Qualitäten (Natale & Ballatore 2020: 5). Eine Besonderheit der Figur der denkenden Maschine liege in der Verwischung der Grenzen zwischen menschlichem Verstehen und statistischer Analyse (Natale & Ballatore 2020: 9), woraus sich spezifische Erwartungen, aber auch Ängste vor der überlegenen Intelligenz von Maschinen ableiteten. Die politikwissenschaftliche Relevanz von Mythen besteht darin, dass sie technische Fähigkeiten überhöhen sowie komplexe, in ihrer weiteren Entwicklung schwer einschätzbare Zusammenhänge vereinfachen und mit einem Eindruck von Unvermeidlichkeit und Geradlinigkeit versehen. Die Entwicklungsoffenheit und die damit verbundenen demokratischen Gestaltungsmöglichkeiten algorithmischer Systeme treten im Rahmen eines mythisch verklärten, linearen Fortschrittsverständnisses in den Hintergrund. Mythen können diese hochpolitische Technologie tendenziell entpolitisieren (Bareis & Katzenbach 2021).

Neben der denkenden Maschine sind es derzeit vor allem die Konzepte der Automatisierung und der Autonomie, die den Diskurs über algorithmische Systeme prägen. Auch wenn Verfahren des maschinellen Lernens tatsächlich auf die Entwicklung sich selbst steuernder Systeme abzielen, handelt es sich dennoch um einen problematischen Mythos, schon weil er

den erheblichen menschlichen Einsatz im Trainieren von Algorithmen und in der Pflege von Datensätzen vernachlässigt (Munn 2022). Wie der nachfolgende Abschnitt zeigt, unterschlägt die Hervorhebung von Automatisierungseffekten aber noch einen weiteren wichtigen Aspekt. In der Gegenüberstellung von demokratischen und algorithmischen Entscheidungsverfahren wird die epistemische, weltdeutende Qualität algorithmischer Systeme erkennbar. Deshalb sollte das Verständnis von KI als Wissenstechnologie mindestens gleichberechtigt neben der Definition als Automatisierungstechnologie stehen. Festzuhalten bleibt jedoch, dass der Begriff der Künstlichen Intelligenz durch seinen Verweis auf die menschliche Intelligenz Erwartungen erzeugt, die die aktuelle Datenanalytik auch in Zukunft nicht einlösen kann.

3. Maschinelles Lernen und demokratische Selbstbestimmung

Die Gegenspieler:innen-Perspektive

In der politikwissenschaftlichen Diskussion spielt das Phänomen der Künstlichen Intelligenz bislang eine randständige Rolle. Im Mittelpunkt der wenigen vorliegenden Beiträge steht die Beziehung zwischen maschinellem Lernen und den Institutionen demokratischer Selbstbestimmung und hier insbesondere der politischen Willensbildung. Die auf den ersten Blick naheliegende Frage lautet, ob maschinelles Lernen sich zu einer Konkurrenz für das demokratische Gemeinwesen entwickeln könnte und wie dieser gegebenenfalls begegnet werden kann. So werfen König und Wenzelburger (2020: 2) die Frage auf, ob algorithmisches Entscheiden womöglich bessere politische Lösungen produziere als die liberale Demokratie und politische Akteure daher zumindest teilweise ersetzen könne. Diese Möglichkeit wird mit dem Verweis auf spezifische Schwächen der repräsentativen Demokratie, die sich derzeit unübersehbar im Umgang mit dem Klimawandel zeigen, konkretisiert: Gemessen an der Datenlage und den durch Modellierungen gewonnenen Prognosen über die erwartbaren Folgen des Klimawandels, „[wirken] Deliberation, demokratische Konkurrenz und die langsame Steuerung über Gesetze [...] hier umständlich und antiquiert“ (Rostalski & Thiel 2021: 62). Am drastischsten formuliert der Historiker Harari (2018) diese Sichtweise. Hätten Demokratien in den Informationsgesellschaften des späten 20. Jahrhundert noch als überlegene Regierungsform gelten können, ändere sich das mit dem Aufkommen des maschinellen Lernens: „Liberal democracy and free-market capitalism see the individual as an autonomous agent constantly making choices about the world [...] It's not so hard to see how AI could one day make better decisions than we do about careers, and perhaps even about relationships. But once we begin to count on AI to decide what to study, where to work, and whom to date or even marry, human life will cease to be a drama of decision making, and our conception of life will need to change.

Democratic elections and free markets might cease to make sense“ (Harari 2018: 12–13).

Um zu verstehen, warum viele wissenschaftliche Disziplinen die Beziehung zwischen Demokratie und maschinellem Lernen als Konkurrenzverhältnis wahrnehmen, hilft es, einen Blick auf ihre Berührungspunkte zu werfen. Aus politikwissenschaftlicher Sicht erklärt sich das unterstellte Spannungsverhältnis zwischen Demokratie und algorithmischen Systemen aus den unterschiedlichen Verfahren der Entscheidungsfindung. Das Versprechen algorithmischer Systeme, so König und Wenzelburger (2020: 3) besteht darin, Aufgaben in einer Weise zu lösen, die menschlichen Fähigkeiten weit überlegen sind. Die kognitiven Aufgaben, um die es dabei geht, bestehen darin, Datensätze mit Blick auf zuvor definierte Zielparameter zu analysieren. Maschinelles Lernen ist, wie es Mackenzie (2017: 7) formuliert, eine historisch spezifische Form der Kalkulation, die auf das Finden und Optimieren mathematischer Funktionen ausgerichtet ist. Ihre Präzision im Bereich der statistischen Verallgemeinerung (Inferenz) liegt jenseits der Möglichkeiten des menschlichen Verstandes (Hildebrandt 2016: 10).

Das zentrale Problem dieser Verfahren für den politischen Prozess bestehe nun darin, so König und Wenzelburger (2020: 5), dass für die Güte politischer Entscheidungen keine vergleichbaren Maßstäbe vorliegen. Im Rahmen einer pluralistisch-liberalen Demokratie könne gerade nicht von einer Einigkeit über gute politische Entscheidungen ausgegangen werden. Im Gegenteil, demokratische Gesellschaften zeichnen sich durch widerstreitende Auffassungen darüber aus, was die richtige politische Entscheidung kennzeichne. Gerade für den Bereich normativer Fragestellungen gelte, dass es sich hier nicht um mathematisch lösbare Entscheidungsprobleme handele, sondern „fundamentally about incommensurable beliefs and values“ (König und Wenzelburger 2020: 5). Legt das maschinelle Lernen den Fokus also auf die Modellierung eines Entscheidungsproblems und die Formulierung berechenbarer Zielparameter, besteht die Herausforderung demokratischer Gesellschaften darin, sich überhaupt erst auf gemeinsame Problembeschreibungen und Zielsetzungen einigen zu können. Der „offene Möglichkeitshorizont“ (Makropoulos 2001) gesellschaftlicher Entwicklung, vor dem demokratische Selbstbestimmung überhaupt erst ihren Sinn erlangt und realisierbar wird, bildet zugleich die epistemische Grundlage der Anfechtbarkeit allen politischen Handelns und des „drama of decision-making“ (Harari 2018).

Im Vergleich zur Vieldeutigkeit gesellschaftlicher Problemlagen und der unterschiedlichen Werthaltungen, mit denen demokratische Prozesse rechnen müssen, stellt sich maschinelles Lernen als reduzierender Modus der Realitätsverarbeitung dar, der wie eine „Gegenspielerin demokratischer Pluralitätsannahme[n] (Koster 2021: 5) wirke. Die mathematischen Verfahren algorithmischer Systeme transportierten „Absolutheitsansprüche“ mitsamt ihrem Glauben an unbestreitbares, objektiv gültiges Wissen in den politischen Raum, um dort auf die demokratische Tradition des diskursiven Ringens um Weltansichten und bessere Argumente zu treffen, so resümiert Koster (2021: 14) die politikwissenschaftliche Sicht auf KI.

Die Gefahr, die aus politikwissenschaftlicher Sicht von algorithmisch ermittelten Handlungsempfehlungen für die Demokratie ausgeht, besteht darin, dass sich das Verständnis des Politischen ändern könnte. Politisches Entscheiden erschiene dann nicht länger als Ergebnis einer Abwägung verschiedener Handlungsoptionen, sondern als Vollzug von mehr oder minder objektiven Notwendigkeiten. Das Bewusstsein für die Unsicherheit und Interpretationsbedürftigkeit allen Wissens würde hierdurch geschwächt. Das darauf referenzierende politische Handeln erschiene deshalb in einem anderen Licht, nämlich als Informationsproblem. Im Szenario einer Zunahme algorithmischer Handlungsvorschläge kommt es entsprechend zu einer Abnahme von Entscheidungen, die sich einer demokratischen Willensbildung und folglich der kollektiven Selbstbestimmung überhaupt noch zurechnen lassen. Deshalb, so Rostalski und Thiel (2021: 59–60), berühre der Einsatz lernender Algorithmen in der Politik die für das demokratische Leben unabdingbare individuelle und kollektive Autonomie, denn „wer für sich selbst nicht länger die Autorenschaft übernimmt, wird auch an den Belangen der Gemeinschaft kein Interesse haben“.

Es ist nicht übertrieben zu sagen, dass die Politikwissenschaft dazu neigt, KI als Gefährdung der Demokratie und ihrer Institutionen zu betrachten. Während politische Regulierungsmaßnahmen im Allgemeinen als Ergebnis institutionalisierter Aushandlungsprozesse verstanden werden, in denen verschiedene Akteure, Interessen und Weltdeutungen aufeinandertreffen und um Diskurs- und Entscheidungshoheit kämpfen, lässt die Diskussion um maschinelles Lernen diese Aushandlungsperspektive fast völlig vermissen. Weil das Verhältnis von KI und Demokratie intuitiv als Konkurrenzverhältnis wahrgenommen werde, so Koster (2021: 6), bleibe die „Analyse des konkreten Aufeinandertreffens beider Logiken hinter einem postulierten Dominanzgewinn der technischen Logik zurück“. Statt empirisch zu prüfen, wie sich die Begegnung von demokratischer Willensbildung und algorithmischen Wahrheitsansprüchen in konkreten gesellschaftlichen Handlungsfeldern

darstelle, ziehe sich die Politikwissenschaft auf eine Bedrohungsperspektive zurück.

Die zunächst plausibel erscheinende Vorstellung einer Rivalität zwischen politischen und algorithmischen Entscheidungslogiken läuft jedoch paradoxerweise Gefahr, ihrerseits einem reduktionistischen Weltbild zu erliegen. Das wäre etwa dann der Fall, wenn die Ausgangsvermutung einer Konkurrenzbeziehung andere Varianten des Zusammenspiels zwischen Demokratie und algorithmischen Systemen systematisch ausblendet. Anstatt sich in einem „Unterwerfungs- oder Übernahmenarrativ“ intellektuell einzurichten, so Koster (2021: 13), sei vielmehr „die Erkundung eines Wechselverhältnisses, vielleicht auch ihres Nebeneinander-Existierens“ wünschenswert. Alles in allem gelte es, die Leistungsfähigkeit von Demokratien nicht zu unterschätzen. Demokratische Öffentlichkeiten seien durchaus in der Lage, algorithmische Systeme kritisch zu reflektieren, und dies vor allem dann und dort, wo diese in gesellschaftliche Handlungs- und Entscheidungszusammenhänge integriert und somit praktisch wirksam werden.

Dieser klugen Beobachtung Kosters kann man hinzufügen, dass wir auch die Wandlungsfähigkeit algorithmischer Systeme nicht unterschätzen sollten. Es gibt keinen Absolutheitsanspruch formulierenden Wesenskern im maschinellen Lernen. Entsprechend könnten lernende Algorithmen auch so entwickelt und eingesetzt werden, dass sie „die Kontingenz politischen Entscheidens deutlich machen“ statt sie zu verbergen (Rolstalski & Thiel 2021: 62). Um die Spielräume algorithmischer Systeme für demokratiefreundliche Nutzungsweisen besser zu verstehen, wird das Verhältnis zwischen Demokratie und maschinellem Lernen im Folgenden aus der Perspektive der Technikentwicklung beleuchtet.

Die Mitspieler:innen-Perspektive

In der Softwareentwicklung gibt es eine lange Tradition der kritischen Auseinandersetzung mit der Arbeitsweise und den Zielsetzungen der eigenen Disziplin. Im Zuge der Expansion des maschinellen Lernens in Wirtschaft und Wissenschaft haben verschiedene Subdisziplinen demokratierelevante Problemstellungen des Datenschutzes, der Diskriminierung verletzlicher gesellschaftlicher Gruppen, der Fairness und der Transparenz zusammengeführt. Das generelle Anliegen von Initiativen wie „Human-Centered AI“ (HCAI) oder „Fairness, Accountability and Transparency in Machine Learning“ (FAT/ML) besteht darin, die Ziele und Gütekriterien in der Entwicklung lernender Systeme so zu erweitern, dass gesellschaftliche Anwendungskontexte und die Lebenslagen der von algorithmischen Entscheidungen betroffenen Menschen schon in der Entstehungsphase Berücksichtigung

finden. Es gelte mit anderen Worten, den „primarily technical worldview“ (Selbst 2018: 2) durch einen „sociotechnical view“ (Selbst 2018: 8) zu ersetzen.

Inzwischen mehren sich die Ansätze, die an der Realisierung einer solchen soziotechnischen Perspektive arbeiten. Zu ihren Zielen gehört, den Nutzer:innen mehr Kontrolle über algorithmische Systeme zu geben (vgl. Shneiderman 2020), die Transparenz und Erklärbarkeit algorithmischer Berechnungen zu erhöhen oder die inhärenten Werthaltungen von Trainings- und Testdatensätzen sichtbar zu machen (Dignum 2022). Statt Datensätze als natürliche Abbildungen der Welt zu akzeptieren, müsse es darum gehen, diese als „situated reflection of a particular worldview“ aufzufassen (Scheuerman et al. 2021: 3). Initiativen wie FAT/ML oder HCAI entstehen in dem Wissen um die Gestaltungsfreiheit und Kontingenz in der Entwicklung algorithmischer Systeme. Ihre Kritik an den vorherrschenden Produktionsnormen manifestiert sich als reflexive Haltung, die einen auch gesellschaftlich wünschenswerten Wandel der eigenen professionellen Normen, Ziele und Praktiken anstrebt (vgl. dazu Selbst 2018).

Verwandte Überlegungen finden sich in der techniksoziologischen Forschung, die ebenfalls die Entwicklungsoffenheit und Kontextabhängigkeit algorithmischer Systeme betont. Hier wird maschinelles Lernen als spezifische Form der Kalkulation (Mackenzie 2015), der Wissensproduktion (Ulbricht 2020) oder als „Beobachtungsformat“ (Heintz 2021) verstanden. Im Verständnis von maschinellem Lernen als Wissenstechnologie drückt sich eine gewisse Akzentverschiebung gegenüber der verbreiteten Gleichstellung von maschinellem Lernen mit der Automatisierung von Entscheidungen (algorithm-based decision making, ADM) aus. Eine Automatisierung von Bewertungs- und Entscheidungsvorgängen finde zwar statt, aber die kritische Algorithmenforschung verortet die gesellschaftliche und politische Relevanz weniger im Vorgang der Automatisierung als in den spezifischen Deutungen und Repräsentationen der Welt, die diesen Automatisierungen vorausgehen. Wie Mackenzie (2015: 444) es formuliert, „the production of prediction is [...] not automatic, although it is being automated“. Das sozialwissenschaftliche Interesse richtet sich entsprechend auf die epistemische, wissens- und wirklichkeitsschaffende Dimension, die den Automatisierungsprozessen innewohnt. Es gehe darum, die „new ways of reading the social world and of intervening in it“ (Fourcade & Johns 2020: 806) auch in ihren performativen, weltverändernden Merkmalen zu verstehen.

Maschinelles Lernen bildet die Welt in Form probabilistischer Modelle ab, die klassifizierende und prädiktive Aussagen herstellen: Das ist wahrscheinlich Spam; dieses Foto zeigt eine Katze. Aber auch: Diese Kandidatin ist vielversprechend und diese Straftäterin wird mit großer Sicherheit (nicht) rückfällig. Die Objektivität der algorithmischen Urteile und Handlungsempfehlungen (und damit auch die kategoriale Andersartigkeit gegenüber dem demokratischen Denken, die dem maschinellen Lernen allgemein zuerkannt wird) entpuppt sich bei näherem Hinsehen allerdings als eine zweifelhafte Zuschreibung, die von den Programmierenden selbst so nicht unbedingt geteilt wird. Wie Dignum (2022: 5) es ausdrückt, stellen algorithmische Systeme eine „accumulation of choices“ dar, also eine Anhäufung von Selektionsentscheidungen unter vielen möglichen Optionen.¹ Automatisierte Entscheidungssysteme kann man sich entsprechend als „manifestation of a proposed solution“ (Bucher 2018: 23) vorstellen. Algorithmen „propose things in and about the world“, so auch Amore (2020: 13), denn in jeder einzelnen Operation eines „apparently autonomous system resides a multiplicity of human and algorithmic judgments, assumptions, thresholds, and probabilities“ (Amore 2020: 64). Anders als der Begriff der „selbstlernenden Algorithmen“ suggeriert, programmieren sich algorithmische Systeme also keineswegs selbst. Gerade in der Formalisierung und Herstellung von Berechenbarkeit liegt der kreative Anteil der Programmierarbeit (Abebe & Barocas 2020: 255).

Ein anschauliches Beispiel dafür bildet „Perspective AI“, ein von Google entwickeltes Analyseprogramm, das lernende Algorithmen mit Datensätzen zur automatischen Identifikation toxischer, das heißt die Diskussionsatmosphäre vergiftender Sprache verknüpft. Wie Rieder und Skop (2021) detailliert beschreiben, ergibt sich die Unterscheidung zwischen Zivilität und Toxizität, die in der algorithmischen Moderation von Onlineforen verwendet wird, keineswegs von selbst. Sie sei kulturell geprägt, immer potenziell diskriminierend und erfordere daher fortlaufende Nachjustierungen. Aufgrund der engen Interaktion zwischen Datensätzen, menschlichen und maschinellen Urteilen liegt es daher nahe, algorithmische Systeme weniger als externe Fremdkörper von Demokratie und Gesellschaft zu verstehen als ein „komplexes Gewebe“ (Rieder und Skop 2021: 12) oder auch als „Mitspieler“ (Koster 2021: 11). Viel relevanter als die bekannte Frage nach der Verteilung von Kontrolle zwischen Menschen und Maschinen, so Rieder und Skop (2021: 12), sei die Beleuchtung der „subtle forms of semantic and moral engineering“, die in der Entwicklung von Kategorien wie sprachlicher Toxizität zum Tragen kommen. Die Botschaft der Autoren lautet, dass wir die künftige gesellschaftliche

¹ Optional ist die Auswahl der Lernalgorithmen, die Definition der Zielvariable, die Zusammenstellung von Trainings- und Testdaten wie auch Optimierungen im Zuge des Trainingsprozesses (vgl. Domingos 2012: 79–80).

Bedeutung des maschinellen Lernens besser erfassen, wenn wir unsere Aufmerksamkeit weniger auf die Unterschiede zwischen algorithmischer Berechnung und selbstbestimmtem Entscheiden richten, sondern mehr auf die Effekte der Interaktion zwischen maschinell und sozialem Handeln.

Zusammenfassend ergeben sich zwei Einsichten für das Zusammenspiel zwischen Demokratie und algorithmischen Systemen. Eine Erkenntnis besteht darin, dass sich die Objektivität der Erkenntnisse und Urteile, vor allem aber die kategoriale Andersartigkeit gegenüber dem demokratischen Diskurs, die dem maschinellen Lernen derzeit gern zuerkannt wird, bei näherem Hinsehen als eine fragwürdige Zuschreibung entpuppt.

Wenn algorithmische Modelle und Berechnungen im Kern nichts anderes als Vorschläge über die zu berechnenden Zusammenhänge darstellen, dann unterscheiden sie sich in dieser Hinsicht nicht grundsätzlich von anderen Formen der Wissensproduktion. Ihre Resultate bleiben den Umständen ihres Zustandekommens (den verwendeten Datensätzen, Berechnungsmethoden und Zielsetzungen) verhaftet, und ihr Geltungsbereich ist offen für politische Anfechtungen. Daraus folgt als eine weitere Einsicht, dass sich die Deutungsmacht algorithmischer Systeme nicht allein ihrer mathematischen Operationsweise verdankt, sondern auch den Kontexten, in denen diese zum Einsatz kommen. Treffend formuliert von Nowotny (2021: 4) gilt: „predictions assume the power of agency that we attribute to them.“ Die Vorstellung einer Rivalität zwischen Demokratie und Künstlicher Intelligenz übersieht die prägende Rolle des Zusammenspiels zwischen algorithmischen Verfahren und Nutzungskontexten.

Die kritische Informatik leitet aus der Gestaltungsoffenheit algorithmischer Systeme praktische Möglichkeiten der politischen Intervention ab: „this creative process of designing a system is a potential opportunity for non-technical stakeholders, including responsible officials and members of the public, to hash out different ways in which a problem might be understood“ (Abebe & Barocas 2020: 254–255). Zwar erforderten algorithmische Systeme das präzise Definieren von Objekten und das Explizieren von Regeln, im Unterschied zu vagen politischen Versprechungen ließen sich solche eindeutigen Formulierungen des Problems und seiner mathematischen Operationalisierung jedoch überprüfen und einer Kritik unterziehen: „Contrary to the claims of many“, so auch (Bryson 2022), „AI and the digital revolution could create a boon for transparency and human understanding“, nämlich dann, wenn der regulatorische Rahmen solche Zielsetzungen fördert und durchsetzt. Tatsächlich hat die

öffentliche Politisierung algorithmischer Systeme in den letzten Jahren Fahrt aufgenommen. Der nächste Abschnitt widmet sich der Frage, welche Aspekte des maschinellen Lernens Gegenstand der Kritik sind und welche Rolle die politische Partizipation hierbei spielt.

4. Algorithmische Systeme als Gegenstand der politischen Partizipation

Maschinelles Lernen ist kein leicht zugänglicher Gegenstand des politischen Handelns. Es müssen hohe Wissenshürden überwunden werden, um konkrete Verfahren in ihrem Einsatzbereich verstehen, bewerten, kritisieren oder gar selbst nutzen zu können. Im Vergleich zum Klimawandel, einer ähnlich anspruchsvollen Materie, bei der viele Jahre der Berichterstattung und öffentlichen Diskussion dazu beigetragen haben, die Komplexität des Themas nach und nach zu reduzieren und so den Weg für ein breiteres gesellschaftliches Interesse und Engagement zu ebnet, befindet sich maschinelles Lernen noch am Anfang einer solchen öffentlichen Lernkurve. Politisches Engagement mit Bezug zu maschinellern Lernen ist aber auch deshalb voraussetzungsvoll, weil der Zugang zum notwendigen Wissen über konkrete Entscheidungssysteme von privatwirtschaftlichen und öffentlichen Akteuren kontrolliert wird, die sich nur ungern in die Karten sehen lassen. Politisches Handeln im Bereich des maschinellen Lernens hat daher mit einer hoch asymmetrischen Informationsverteilung zu kämpfen. Diesen Schwierigkeiten steht die große gesellschaftliche und wirtschaftliche Relevanz gegenüber, aber auch die häufig unterschätzten Gestaltungsmöglichkeiten und -bedarfe dieser neuen Querschnittstechnologie.

Die Praxis der politischen Partizipation hat sich in den letzten Jahrzehnten so stark diversifiziert, dass die Abgrenzung zu nichtpolitischem Handeln zunehmend schwerfällt. Stand bis in die 1960er-Jahre die Teilnahme an Wahlen im Zentrum der Aufmerksamkeit, erweiterte sich das Spektrum politischer Artikulationsweisen seither immer weiter und scheint sich spätestens seit der Digitalisierung des Engagements einer klaren Abgrenzung zu entziehen. Im Kern handelt es sich bei politischer Partizipation um freiwillige Handlungen (oder Boykotte) nichtprofessionell engagierter Akteure, die sich auf kollektive gesellschaftliche Anliegen richten (vgl. dazu Dritter Engagementbericht 2020; van Deth 2016). Der Intention nach zielt politische Partizipation auf die Problematisierung und Politisierung von Sachverhalten. „Polarisierung“ bedeutet, dass Themen in den Raum politischen Deliberierens und Entscheidens hineingetragen werden und/oder dass ihre Entscheidbarkeit und Kontingenz, nämlich „the capacity of things to be different“ (Hay 2007: 65–66) sichtbar und praktisch erfahrbar gemacht wird (Haunss & Hofmann 2015). Eine solch weite Definition politischen Engagements ist im

Kontext von KI nicht zuletzt deshalb sinnvoll, weil viele politische Interventionen außerhalb der Sphäre des Regierungshandelns, etwa im Umfeld der Entwicklung und Nutzung von KI, oder auch in sich institutionalisierenden Zwischenräumen von akademischer Forschung und Zivilgesellschaft angesiedelt sind.

Ein auffälliges Merkmal der Politisierung von KI zeigt sich gegenwärtig in der sektorübergreifenden Organisation. Vielfach kommt es zur Zusammenarbeit von Nichtregierungsorganisationen (NGOs), Stiftungen, Datenschutzbehörden und wissenschaftlicher Forschung, aber auch zwischen universitären und wirtschaftlichen „machine learners“ sowie zu gemeinsamen Projekten von Zivilgesellschaft und klassischen Massenmedien (vgl. für einen europäischen Überblick, AlgorithmWatch 2020). Der zunehmende Einsatz automatisierter Entscheidungssysteme hat in den letzten Jahren nicht nur bestehende netzpolitische Organisationen auf den Plan gerufen, sondern auch zu einer Reihe von Neugründungen geführt: NGOs wie dem Ada Lovelace Institute (2018, Großbritannien), AlgorithmWatch (2017, Deutschland), AI Now (2017, USA) oder Data and Society (2014, USA) ist eine ähnliche Agenda gemeinsam. Betont wird, dass man sich nicht gegen den Einsatz maschinellen Lernens per se wende, sondern dafür kämpfe, dass dieser im Einklang mit demokratischen Grundsätzen erfolgt: „The message for policy-makers couldn't be clearer. If we truly want to make the most of their potential, while at the same time respecting human rights and democracy, the time to step up, make those systems transparent, and put ADM wrongs right, is now“ (AlgorithmWatch 2020).

Den größten Raum im politischen Engagement nimmt gegenwärtig die Problematisierung automatisierter Entscheidungssysteme ein. Zivilgesellschaftliche Initiativen haben in den letzten Jahren Diskriminierungen durch algorithmische Entscheidungssysteme in Staat und Wirtschaft skandalisiert, die Offenlegung des Quellcodes dieser Systeme gefordert, aber auch in pragmatischer Absicht konkrete Instrumente und Verfahren entwickelt, um dem Ziel einer gesellschaftlich wünschenswerten Integration lernender Algorithmen in den Alltag näherzukommen. Dazu gehören Methoden zur Folgenabschätzung algorithmischer Systeme, die Entwicklung von Leitlinien für Audits und Zertifizierungen, Experimente mit alternativen Empfehlungsalgorithmen (vgl. dazu Edwards & Veale 2017) und Datenspenden. Wichtig sind darüber hinaus Kampagnen für das Verbot von menschenrechtsgefährdenden algorithmischen Systemen (Hartzog 2018), die Beteiligung an gesetzlichen Regulierungsvorhaben sowie der Kampf um eine effektive politische Aufsicht, das heißt mehr Transparenz, Rechenschaftspflichtigkeit und institutionalisierte gesellschaftliche Partizipation. Im Folgenden

werden zwei besonders relevante Interventionsformen näher betrachtet, nämlich die Kritik an algorithmenbasierten Entscheidungssystemen und die Beteiligung an staatlichen Regulierungsverfahren.

Einsatz algorithmenbasierter Entscheidungssysteme

In vielen Ländern reagieren zivilgesellschaftliche Initiativen auf den inzwischen schnell um sich greifenden Einsatz lernender Algorithmen in der Wirtschaft und den öffentlichen Verwaltungen. So hat AlgorithmWatch in den Jahren 2019 und 2020 zusammen mit der Bertelsmann Stiftung Länderstudien über gelungene, vor allem aber weniger gelungene Einsätze von algorithmischen Entscheidungssystemen in der EU erstellt. Eine leitende Frage dabei war, „wie sich Automatisierung auf unsere Institutionen, unsere Normen und Werte, ja auf die Demokratie insgesamt auswirkt“ (AlgorithmWatch 2020: 6). Die dokumentierten Fälle reichen von der vorhersagenden Polizeiarbeit (predictive policing) und der automatisierten Zulassung von Studierenden über den Einsatz von Gesichtserkennungsprogrammen bis zur Prüfung von Sozialleistungsansprüchen und der Aufdeckung von Steuerbetrug.

Obwohl der Bericht ausdrücklich auch positive, vorbildhafte Nutzungsweisen in den Blick nehmen wollte, dominieren doch die Fälle, in denen automatisierte Entscheidungsverfahren die Grundrechte betroffener Bürger:innen unterlaufen. Ein zentraler Kritikpunkt des Berichts lautet, dass die Automatisierungsverfahren in der Regel ohne öffentliche Diskussion eingeführt werden, ihre Entscheidungslogik verheimlicht wird und die konkrete Zielsetzung für Außenstehende (und womöglich auch für die Verwaltungen selbst) häufig unklar bleibt.² Sofern Aufsichtsregelungen bestehen, werden sie offenbar häufig ignoriert oder aber ihre Umsetzung hält mit dem Einsatz algorithmischer Systeme nicht Schritt (AlgorithmWatch 2020: 10–11; vgl. auch Ada Lovelace Institute 2021). Angesichts dieser Situation eignet sich zivilgesellschaftliches Engagement eine Art informelle Aufsichtsfunktion an, die darin besteht, Evidenz für Diskriminierungen und Fehlentscheidungen zu sammeln und gesellschaftlichen Druck zugunsten einer Offenlegung der Entscheidungssysteme sowie einer effektiveren Kontrolle aufzubauen. Diese zivilgesellschaftliche Aufsichtsfunktion sieht sich allerdings mit vielen Widerständen konfrontiert.

² Exemplarisch hierfür das Scheitern der algorithmischen Bewertung von Schulabschlüssen in Großbritannien (Jones & Safak 2020).

Weil private wie auch öffentliche Anwenderorganisationen den Zugang zu algorithmischen Verfahren und Datensätzen zumeist verweigern, richtet sich das Repertoire zivilgesellschaftlicher Akteure notgedrungen auf die Entwicklung von Strategien zur Bekämpfung der bestehenden Informationsasymmetrie aus. Bezogen auf die öffentliche Hand stellen Informationsfreiheitsanfragen ein wichtiges Instrument dar, die in zunehmendem Umfang durch Gerichtsverfahren ergänzt werden. Die Erfolgsaussichten solcher Klagen sind allerdings unsicher.³ Da gegenüber der Privatwirtschaft nur eingeschränkt Ansprüche auf Offenlegung von Algorithmen geltend gemacht werden können, müssen hier andere Wege zur Herstellung von Transparenz und Rechenschaftspflichtigkeit eingeschlagen werden.⁴ Um etwa die Logik algorithmischer Nachrichtenkuratierung oder die Ausspielung von Werbung auf digitalen Plattformen nachvollziehen zu können, verwenden NGOs Crowdsourcing-Methoden. Mithilfe von Browser-Add-ons können individuelle Nutzungshistorien durch Datenspenden gesammelt und mit Reverse Engineering ausgewertet werden, um Einblicke in die Logik von Empfehlungsalgorithmen oder den Umfang der Personalisierung von Informationsströmen zu gewinnen (vgl. dazu auch Rieder & Hofmann 2020).

AlgorithmWatch experimentiert seit mehreren Jahren mit dem Instrument der Datenspende, um die transparenzschaffenden Möglichkeiten und Grenzen dieses Verfahrens genauer zu erkunden. So sollte das 2018 zusammen mit der Open Knowledge Foundation begonnene Projekt „OpenSCHUFA“ mehr Licht ins Dunkel der opaken Kreditwürdigkeitsberechnung der größten deutschen Auskunftei bringen.⁵ Die hierfür lancierte Finanzierungs- und Datenspendenkampagne konnte die Erwartungen jedoch nicht erfüllen.⁶ Ein weiteres Datenspendenprojekt zur Untersuchung der Empfehlungsalgorithmen von YouTube im Vorfeld der Bundestagswahl im Jahr 2021 hatte ebenfalls mit Problemen zu kämpfen (vgl. Kahlert 2022). Nicht überraschend ist, dass sich die Privatwirtschaft gegen Versuche externer Kontrolle wehrt. Ihre Nutzungsbedingungen schließen automatische Analysen von Nutzungsdaten häufig aus; (nicht nur) Facebook ist mehrfach gegen externe Analysen seiner

³ In den Niederlanden konnte der Einsatz von SyRI, einem System zur Aufdeckung des Missbrauchs von Sozialleistungen unterbunden werden, nachdem ein Gericht urteilte, dass SyRI gegen die Europäischen Menschenrechtskonvention verstößt (Parbel 2020). In Spanien scheiterte dagegen der Versuch, Zugang zum Quellcode eines fehlerhaften Systems zur Prüfung der Anspruchsberechtigung für Stromkostenzuschüsse zu erlangen. Nachdem die spanische Regierung dessen Offenlegung verweigerte, unterlag die Fundación Civio Anfang 2022 nun auch vor dem Verwaltungsgericht (vgl. Valdivia & de la Cueva 2022).

⁴ Aktuelle Regulierungsansätze wie der deutsche Medienstaatsvertrag oder das EU-Gesetz über digitale Dienste schaffen neue Transparenzaufgaben für Plattformen und soziale Netzwerke, die aber viele Fragen offenlassen.

⁵ Die Auswertung von rund 2.000 Datensätzen durch den Bayerischen Rundfunk und Spiegel ergab, dass die Auskunftei in einigen Fällen negative Bewertungen vornimmt, die nicht durch die jeweilige Kredithistorie begründet sind, und dass bei einem Viertel der datenspendenden Personen zu wenig Daten vorlagen, um Prognosen über mögliche Zahlungsausfälle überhaupt erstellen zu können (SPIEGEL Data & BR Data 2018).

⁶ Es beteiligten sich nur wenige Tausend Menschen an der Kampagne, und es gelang nicht, diejenigen Bevölkerungskreise zu erreichen, die von fehlerhaften Scores mutmaßlich am stärksten betroffen sind.

Kuratierungsprinzipien vorgegangen (Kayser-Bril 2021). Aufgrund dieser manifesten Abwehrstrategien zielt die politische Auseinandersetzung über einzelne problematische Fälle hinaus auf die Etablierung effektiver Transparenzregeln (Keller & Drake 2022).

Regulierung algorithmischer Systeme

Ein weiteres wichtiges zivilgesellschaftliches Betätigungsfeld bilden Stellungnahmen zu staatlichen Regulierungsvorhaben. Der Entwurf der Europäischen Kommission für ein Gesetz über Künstliche Intelligenz (Artificial Intelligence Act, AIA) ist derzeit das wichtigste Thema im europäischen Raum.⁷ Ende 2021 haben 115 Organisationen einen Aufruf veröffentlicht, der fordert, die Wahrung der Menschenrechte in den Mittelpunkt der Regulierung von Künstlicher Intelligenz zu stellen. Darüber hinaus kritisiert der Aufruf den risikobasierten Ansatz des AIA-Vorhabens (EDRi et al. 2021). In einer im Frühjahr 2022 vorgelegten Expertise für das britische Ada Lovelace Institute konkretisiert Edwards (2022) die Gründe für die Ablehnung des Gesetzentwurfs. So beziehe der Entwurf seine „Inspiration“ aus den Vorschriften zur Produktsicherheit (product safety legislation) und behandle daher die Anbieter von AI-Systemen analog zu den Herstellern herkömmlicher Produkte wie Spülmaschinen oder Spielzeug. Aber: „AI is not a dishwasher and the way downstream deployers use it and adapt it, may be as significant as how it is originally built“ (Edwards 2022: 6). Weil algorithmenbasierte Systeme häufig das arbeitsteilige Resultat vieler Anbietenden seien und sich zudem mit ihren Anwendungskontexten gravierend ändern können, sei die Vorgehensweise des AIA, nämlich „holistically evaluating the risk of such a system in the abstract [...]“ (Edwards 2020: 6), nicht möglich .

In den kritischen Stellungnahmen zivilgesellschaftlicher Organisationen zeigen sich spezifische Stärken und Schwächen der sektorübergreifenden Politisierung maschinellen Lernens. Die Verknüpfung akademischer Expertise mit gesellschaftlichem Engagement zum Schutz von Menschenrechten und zur Bekämpfung von Diskriminierung sorgt für eine hohe Anschlussfähigkeit politischer Interventionen an die Denk- und Arbeitsweise legislativer und exekutiver Regulierungsinstanzen wie auch für eine ausgewiesene Fachkompetenz über einzelne Regulierungsinstrumente. Deutlich wird dies etwa in Bezug auf den risikobasierten Ansatz des AIA. Die Stellungnahmen zum AIA-Entwurf sind auch deshalb bemerkenswert, weil sie das öffentliche Verständnis von Künstlicher Intelligenz politisieren. Der Diskurs über die (supra-)staatliche Regulierung liest sich wie eine Form von „ontological politics“ (Mol

⁷ COM/2021/206 final, online verfügbar unter: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>.

1999), weil hier um eine angemessene kollektive Interpretation des Wesens lernender technischer Systeme gerungen wird: Wenn sich die Anwendungsrisiken algorithmischer Systeme nicht analog zu jenen von Spülmaschinen konzipieren lassen, an welchen Vorbildern sollten sich Regulierungsmaßnahmen stattdessen orientieren? Das Ziel der staatlichen Einhegung, so zeigt sich an dieser Diskussion auch, zementiert spezifische Definitionen von KI. Die können allerdings im Zuge ihrer Politisierung eine kritische Reflexion durchlaufen (vgl. Bryson 2022).

Selektivität politischer Partizipation

Obwohl das politische Engagement viele Facetten und Entwicklungsstadien algorithmischer Systeme adressiert, ist es doch selektiv. Gemeinsam ist den einzelnen Initiativen der strikte Bezug zur Gegenwart. Im Zentrum der Kritik stehen ganz überwiegend Verfahrensfragen, konkrete Normverstöße oder die Manipulation von Bürger:innen und Öffentlichkeit. Nur am Rande aufgegriffen wird dagegen die epistemische, weltdeutende Dimension algorithmischer Systeme. Zwar betont die kritische Informatik die Gestaltungsoffenheit und Kreativität in der Entwicklung lernender Algorithmen, aber solche Einsichten übersetzen sich nur schwer in die politische Strategiebildung zivilgesellschaftlicher Organisationen. Ähnliches gilt für die Konzeption algorithmischer Systeme als „Mitspieler“ in spezifischen Anwendungskontexten. Der politische Fokus auf automatisierte Entscheidungssysteme verliert gelegentlich aus dem Blick, dass die Ursachen für Diskriminierungen häufig nicht in den algorithmischen Entscheidungssystemen selbst, sondern in den Logiken und Strukturen ihrer Anwendungskontexte liegen (Eubanks 2018: 152; Hoffmann 2019). Um automatisierte Entscheidungssysteme als interaktive Mitspieler politisch kritisieren zu können, bedürfte es jedoch des Zugangs zu den jeweiligen Einsatzorten, der zivilgesellschaftlichen Akteur:innen jedoch zumeist verwehrt wird.

Im Ergebnis trägt dieser Fokus politischer Partizipation dazu bei, dass gegen den Einsatz algorithmischer Systeme vor allem dann protestiert wird, wenn die darin enthaltenen Welt- und Menschenbilder belegbare Diskriminierungen hervorrufen. Die langfristigen gesellschaftlichen, politischen und semantischen Transformationsprozesse im Zuge von Datafizierung und Algorithmisierung sind zwar zunehmend beliebte Forschungsfragen in den Sozialwissenschaften, im politischen Handeln aber erlangen sie faktisch keine Aufmerksamkeit. Das epistemische Programm, das den Einsatz algorithmischer Systeme als Wissenstechnologie vorantreibt, bleibt politisch ein blinder Fleck. Ein Grund dafür könnte darin bestehen, dass die wirklichkeitsstiftende und transformative Macht von Algorithmen zu

abstrakt ist, um sich für politische Kampagnen zu eignen; ein anderer, dass externe Finanzierungen, etwa durch Stiftungen, leichter für greifbare, gegenwartsbezogene Projekte zu gewinnen sind als für grundlegendere Fragestellungen des demokratischen Wandels. Die Gefahr dieses blinden Flecks liegt darin, dass sich das Zeitfenster für eine Politisierung der epistemischen Macht des maschinellen Lernens schließen könnte. Zu erwarten ist, dass algorithmische Verfahren im Zuge ihrer Nutzung zunehmend normalisiert werden und die öffentliche Sensibilität für ihre Deutungs- und Nutzungsoffenheit abnimmt.

Ein weiteres Merkmal des praxisnahen Horizonts politischen Engagements zeigt sich in einem Mangel an echten Gegenentwürfen zum allgemeinen Diskurs über Künstliche Intelligenz. Es ist nicht erkennbar, wie man sich die Entwicklung und Nutzung maschinellen Lernens zum Wohle aller jenseits von manipulierenden Empfehlungsalgorithmen und Entscheidungsautomatisierung vorzustellen hätte. Es fehlt an überzeugenden Visionen einer „good AI society“ (Cath et al. 2018), die die große Bedeutung dieser Technologien einerseits und die darin enthaltenen Wahlmöglichkeiten andererseits so ausbuchstabiert, dass die darin liegenden Herausforderungen demokratisch bearbeitbar würden. Der Mangel an richtungsweisenden Gegenentwürfen und das gleichzeitige Fehlen eines kollektiven Nachdenkens über die langfristigen Transformationsprozesse im Gefolge dieser neuen Querschnittstechnologie könnten darauf hinauslaufen, dass gesellschaftliche Gestaltungsmöglichkeiten unausgeschöpft und der demokratischen Selbstbestimmung entzogen bleiben.

5. Exkurs: Demokratischer Wandel im Schatten des maschinellen Lernens

Der folgende spekulative Exkurs geht von der Vermutung aus, dass man der künftigen Bedeutung algorithmischer Systeme für die Demokratie näherkommt, wenn man diese nicht als Gegenspieler:innen, sondern als Mitspieler:innen betrachtet (Koster 2022; Rieder & Skop 2021). Die Ausbreitung des maschinellen Lernens erfolgt ja nicht in Form einer feindlichen Übernahme, sondern sie setzt voraus, dass der algorithmischen Berechenbarkeit politisch ein spezifischer Mehrwert zugeschrieben wird. Folglich geht es darum, Anschlusspunkte zu identifizieren, an denen sich demokratisches Denken und Handeln offen für die spezifischen Ressourcen des maschinellen Lernens zeigt. Aus der heutigen Perspektive zeichnen sich solche Punkte im Nachdenken über die praktischen Probleme der Demokratie, insbesondere aber auch im Bereich ihrer Regulierungsfelder und -instrumente ab.

Eine wichtige Ressource des maschinellen Lernens besteht in Datenanalysen, die als neues „Beobachtungsformat“ (Heintz 2021) gelten können. So wie die Bevölkerungsstatistik eine zentrale Rolle in der Entdeckung der Gesellschaft als einem „eigenständigen Realitätsbereich“ gespielt hat (Heintz 2021: 149), so könnten große Datensätze dazu beitragen, dass sich die Gesellschafts- und Menschenbilder ändern, die im Gefolge der Statistik entstanden sind und die das kollektive Selbstbild bis heute prägen. Als Wissenstechnologien verstanden, fungieren Daten und lernende Algorithmen nicht nur als analytische Werkzeuge, sondern immer auch als Medium, durch das sich eine berechenbarer werdende Welt neu darstellt und für andere Repräsentations- und Interpretationsweisen öffnet (Panagia 2020: 3).

Die Auswertung psychosozialer Daten setzt gegenwärtig vor allem das demokratische Leitbild des autonom handelnden Subjekts unter Rechtfertigungsdruck (vgl. für einen Überblick Block & Dickel 2020). Der Tradition des humanistischen Denkens folgend sind Menschen in der Lage, ihr Leben individuell wie kollektiv unter Gebrauch ihrer Vernunft selbstbestimmt zu regeln. Dieser Vorstellung des rational handelnden Subjekts widersprechen unter anderem die Verhaltensökonomik und -psychologie, die mithilfe empirischer Analysen belegen möchten, dass menschliches Entscheiden gerade nicht vernünftig ist, sondern im Alltag sogenannten „Urteilsheuristiken“ folgt, also unbewussten Handlungsregeln, die – nicht überraschend – häufig suboptimal ausfallen (Kahnemann 2003).

Beispiele für die beschränkte Rationalität individuellen und kollektiven Handelns gibt es in der Geschichte wie der Gegenwart natürlich unendlich viele; menschengemachte Katastrophen wie Klimawandel oder Kriege zeugen davon. Die entscheidende Frage in diesem Zusammenhang ist, ob wir unter dem Eindruck zunehmender algorithmischer Prognosen und Handlungsempfehlungen zu einer Neubewertung der unübersehbaren Diskrepanz zwischen dem Ideal des rational handelnden Menschen und der suboptimalen Realität gelangen. Zugespitzt: Erweisen wir uns weiterhin als enttäuschungsfest und verteidigen die Möglichkeit demokratischer Selbstbestimmung oder schließen wir uns der (von der Verhaltensökonomik vertretenen) Überzeugung an, dass wir algorithmischen Maschinen zu unserem eigenen Vorteil ein größeres Mitspracherecht einräumen sollten? Zur Diskussion steht, wie es Hayles (2005: 132) formuliert, nichts weniger als unser Verständnis des menschlichen Denkens, Handelns und Wahrnehmens, „how we understand what it means to be human“ (vgl. dazu auch Hildebrandt 2016).

Im Zusammenhang mit den verhaltensökonomischen Annahmen über die beschränkte Rationalität menschlichen Handelns entstehen neue Steuerungsparadigmen und Regulierungsinstrumente, deren Pionier unter dem Begriff „Nudging“ bekannt geworden ist (Thaler & Sunstein 2008). Dabei handelt es sich um niedrigschwellige, individualisierte Maßnahmen, die Anstöße für Verhaltensänderungen und/oder den Erwerb neuer Entscheidungskompetenzen bieten. Ihre Operationsweise lässt sich an der Bekämpfung von Falschinformationen im Internet verdeutlichen. Der Konsum und die Verbreitung von Desinformation gelten gegenwärtig als Gefahr für das demokratische Gemeinwesen und werfen daher die Frage nach adäquaten politischen Antworten auf (Bennett & Livingston 2018). Die Verhaltensökonomie experimentiert dazu mit verschiedenen Varianten des „educational nudging“, die die Kompetenzen der Bürger:innen darin stärken sollen, die Qualität von Informationen besser einzuschätzen. Zu diesem Zweck werden „epistemic cues“ in die unmittelbare Handlungsumgebung, die sogenannten „Entscheidungsarchitekturen“ von Menschen eingeschleust (Lorenz-Spreen et al. 2020: 1103). Eine andere Version der „Stupser“ beabsichtigt, die überwiegend unbewusst wirkende „Steuerungsmacht sozialer Normen“ gegen die Verbreitung von Desinformation zu mobilisieren (Lorenz-Spreen et al. 2020: 1105–1106). Hierfür erhalten die Zielpersonen Hinweise darüber, wie Dritte zuvor auf einen Artikel oder ein Video reagiert haben.

Im Begriff der „Entscheidungsarchitektur“ drückt sich die Erwartung aus, dass sich das individuelle Verhalten mithilfe von Datenanalysen vorhersagen und durch die strategische Veränderung der jeweiligen Rahmenbedingungen zielgenau beeinflussen lässt (Thaler et al. 2012). Digitale Plattformen wie Facebook oder Amazon perfektionieren diese Steuerungslogik schon seit Jahren, indem sie das Verhalten der Nutzer:innen fortlaufend beobachten und durch Anpassungen der digitalen Handlungsumgebung zu kontrollieren versuchen (vgl. Stark 2018). Im Unterschied zur Plattformökonomie verfolgt die algorithmische Verhaltensregulierung dem eigenen Verständnis nach durchaus emanzipatorische, demokratieförderliche Ziele. So soll die beschränkte „informationelle Autonomie“ des Einzelnen unterstützt werden, damit alle Menschen als „choice architect“, das heißt als Strateg:innen ihrer Entscheidungsumgebung, agieren können (Lorenz-Spreen et al. 2020: 1106).

Die systematische Erhebung und Auswertung von Datenspuren des Verhaltens, der Kontaktnetzwerke, aber auch körperlicher Merkmale wie Pulsschlag, Atem- oder Schlafrhythmus, erzeugen eine bislang unbekannte Tiefenschärfe in der gesellschaftlichen Beobachtung, die gleichzeitig hoch detailliert, global ausgreifend und echtzeitlich ist. Ulbricht (2020) hat für die politische Nutzung dieser Daten den Begriff des „demos scraping“, im wörtlichen Sinne also das Zusammenkratzen von Informationen über das Staatsvolk vorgeschlagen. Das Beispiel der Verhaltensinterventionen gibt eine Vorahnung davon, dass und wie algorithmische Regulierungsinstrumente das Repertoire staatlicher (aber auch privater) Regulierung erweitern. Es entstehen Möglichkeiten der mikroskopisch genauen Intervention, die sogar unterhalb der Bewusstseinsschwelle des Demos ansetzen können.

Politikwissenschaftlich gesehen dehnt sich mit dem maschinellen Lernen der Raum des politisch Verfügbaren und Entscheidbaren aus.⁸ Persönlichkeitstypen, Gefühle und Impulse lassen sich nun als Muster erfassen und auf neue Weise politisch bearbeiten. Mit der Verbreitung prädiktiver Analysen rücken auch die hypothetischen Folgen des Handelns zunehmend in den Bereich des Erkenn- und Beeinflussbaren. Wie die Prognosen zum Klimawandel demonstrieren, lassen sich die langfristigen Effekte politischen Handelns bzw. Nichthandelns immer präziser kalkulieren, sodass sich auch der Raum des politisch zu Verantwortenden immer weiter in Richtung einer probabilistischen Zukunft aufspannen dürfte. Anders formuliert scheint es nicht vollkommen abwegig, dass wir individuell wie auch kollektiv gegenüber einer algorithmisch erzeugten Wirklichkeit zumindest dann begründungspflichtig werden, wenn wir von ihren Empfehlungen und Nudges abweichen. Darauf jedenfalls läuft die Vermutung von Harari (2018: 13) hinaus, der zufolge algorithmische Systeme womöglich irgendwann die besseren Entscheidungen treffen: „Imagine Anna Karenina taking out her smartphone and asking Siri whether she should stay married to Karenin or elope with the dashing Count Vronsky.“ Die Hoffnung, fatale Fehlentscheidungen mithilfe von Algorithmen vermeiden zu können, mag sich tatsächlich als eine schwer zu widerstehende Versuchung erweisen.

Das politische Dilemma, das Yuval Harari vor seinem Publikum entfaltet, liegt in den Kosten, die demokratische Gesellschaften für eine probabilistische Kontrolle über ihre individuelle und kollektive Zukunft zu tragen hätten. Der Preis für eine allgemeine Anerkennung algorithmischer Urteile bestünde in der Ausbreitung von Tyrannei, nämlich einer Verschiebung

⁸ Hartmut Rosa (2020: 21 ff.) beschreibt das Verfügbarmachen der Welt anhand von vier Dimensionen: Erkennbarkeit, Erreichbarkeit, Kontrollierbarkeit und Nutzbarkeit.

politisch legitimerter Autorität zu vernetzten Maschinen, der Konzentration datenbasierter Macht in wenigen Händen und in dem allmählichen Verblässen von Autonomie als demokratischer Norm, einer unverzichtbaren Grundvoraussetzung demokratischen Zusammenlebens.

Obwohl es eher wahrscheinlich ist, dass sich unser Verständnis von Autonomie im Laufe der gesellschaftlichen Integration algorithmischer Wissenstechnologien verändert, folgt daraus natürlich nicht zwangsläufig eine Unterwanderung der Demokratie. Im Sinne gesellschaftlicher Selbstbestimmung wird es darauf ankommen, ein realistischeres Bild nicht nur von den Möglichkeiten, sondern vor allem auch von den Grenzen und Schwächen maschinellen Lernens zu entwickeln. Dazu gehört, sich vom Mythos der „denkenden Maschine“ zu verabschieden und algorithmische Urteile als spezifische Lesarten neben anderen Lesarten der Welt einzuordnen. Dann kann maschinelles Lernen den politischen Diskurs tatsächlich erweitern und bereichern, statt ihn überflüssig zu machen.

6. Literatur

- Abebe, R., Barocas, S., Kleinberg, J., Levy, K., Raghavan, M., & Robinson, D. G. (2020). Roles for Computing in Social Change. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 252–260. <https://doi.org/10.1145/3351095.3372871>
- Ada Lovelace Institute, AI Now Institute and Open Government Partnership (2021). *Algorithmic accountability for the public sector. Learning from the first wave of policy implementation*. <https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector/>
- AlgorithmWatch & Bertelsmann Stiftung (2020). *Automating Society Report 2020*. <https://automatingsociety.algorithmwatch.org/wp-content/uploads/2020/12/Automating-Society-Report-2020.pdf>
- Amoore, L. (2020). *Cloud Ethics – Algorithms and the Attributes of Ourselves and Others*. <https://www.dukeupress.edu/cloud-ethics>
- Bareis, J., & Katzenbach, C. (2021). Talking AI into Being: The Narratives and Imaginaries of National AI Strategies and Their Performative Politics. *Science, Technology, & Human Values*, 1–27. <https://doi.org/10.1177/01622439211030007>
- Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, 33(2), 122–139. <https://doi.org/10.1177/0267323118760317>
- Block, K., & Dickel, S. (2020). Jenseits der Autonomie: Die De/Problematisierung des Subjekts in Zeiten der Digitalisierung. *BEHEMOTH – A Journal on Civilisation*, 13(1), 109–131. <https://doi.org/10.6094/behemoth.2020.13.1.1040>
- Bryson, J. J. (2. März 2022). Europe Is in Danger of Using the Wrong Definition of AI. *Wired*. <https://www.wired.com/story/artificial-intelligence-regulation-european-union/>
- Bucher, T. (2018). *If ... Then: Algorithmic Power and Politics*. <https://www.sack.de/bucher-ifthen-algorithmic-power-and-politics/9780190493035>
- Cardon, D., Cointet, J.-P., Mazières, A., & Libbrecht, E. (2018). Neurons spike back: The invention of inductive machines and the artificial intelligence controversy. *Rezeaux*, 211(5), 173–220.
- Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial Intelligence and the „Good Society“: The US, EU, and UK approach. *Science and Engineering Ethics*, 24(2), 505–528. <https://doi.org/10.1007/s11948-017-9901-7>
- Dignum, V. (2022). *Relational Artificial Intelligence*. <https://doi.org/10.48550/arXiv.2202.07446>
- Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 78–87. <https://doi.org/10.1145/2347736.2347755>
- Dritter Engagementbericht: Zukunft Zivilgesellschaft: Junges Engagement im digitalen Zeitalter. (2020). BMFSFJ. <https://www.bmfsfj.de/bmfsfj/service/publikationen/dritter-engagementbericht-156434>

- EDRi, Bits of Freedom, Fair Trials, accessnow, & Algorithm Watch. (2021). *An EU Artificial Intelligence Act for Fundamental Rights. A Civil Society Statement*. <https://edri.org/wp-content/uploads/2021/12/Political-statement-on-AI-Act.pdf>
- Edwards, L. (2022). *Regulating AI in Europe: Four problems and four solutions*. Ada Lovelace Institute. <https://www.adalovelaceinstitute.org/report/regulating-ai-in-europe/>
- Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St Martin's Press.
- Fourcade, M., & Johns, F. (2020). Loops, ladders and links: The recursivity of social and machine learning. *Theory and Society*, 49(5), 803–832. <https://doi.org/10.1007/s11186-020-09409-x>
- Harari, N. (2018). Why Technology Favors Tyranny. *The Atlantic*. <https://www.theatlantic.com/history/>
- Hartzog, W. (2. August 2018). Facial Recognition Is the Perfect Tool for Oppression. *Medium*. <https://medium.com/s/story/facial-recognition-is-the-perfect-tool-for-oppression-bc2a08f0fe66>
- Haunss, S., & Hofmann, J. (2015). Entstehung von Politikfeldern – Bedingungen einer Anomalie. *der moderne staat – Zeitschrift für Public Policy, Recht und Management*, 8(1), 29–49. <https://doi.org/10.3224/dms.v8i1.19109>
- Hay, C. (2007). Politics, Participation and Polarization. *Why We Hate Politics* (S. 60–177). Polity Press.
- Hayles, K. (2005). Computing the Human. *Theory Culture & Society*, 22, 131–151. <https://doi.org/10.1177/0263276405048438>
- Heintz, B. (2021). Big Observation – Ein Vergleich moderner Beobachtungsformate am Beispiel von amtlicher Statistik und Recommendersystemen. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 73(1), 137–167. <https://doi.org/10.1007/s11577-021-00744-0>
- Hildebrandt, M. (2016). Law as Information in the Era of Data-Driven Agency. *The Modern Law Review*, 79(1), 1–30. <https://doi.org/10.1111/1468-2230.12165>
- Hildebrandt, M. (2020). The Artificial Intelligence of European Union Law. *German Law Journal*, 21, 74–79. <https://doi.org/10.1017/glj.2019.99>
- Hoffmann, A. L. (2019). Where fairness fails: Data, algorithms, and the limits of antidiscrimination discourse. *Information, Communication & Society*, 22(7), 900–915. <https://doi.org/10.1080/1369118X.2019.1573912>
- Jasanoff, S., & Sang-Hyun, K. (2015). *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*. University of Chicago Press. <https://press.uchicago.edu/ucp/books/book/chicago/D/bo20836025.html>
- Jordan, M. I. (2019). Artificial Intelligence – The Revolution Hasn't Happened Yet. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.f06c6e61>
- Joyce, K., Smith-Doerr, L., Alegria, S., Bell, S., Cruz, T., Hoffman, S. G., Noble, S. U., & Shestakofsky, B. (2021). Toward a Sociology of Artificial Intelligence: A Call for Research on Inequalities and Structural Change. *Socius*, 7. <https://doi.org/10.1177/2378023121999581>

- Kahler, P. (28. Februar 2022). YouTubes Wahl. *Medium*. <https://medium.com/sts-ens/youtubes-wahl-ea9c4df1297e>
- Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. *The American Economic Review*, 93(5), 1449–1475.
- Kayser-Bril, N. (o. J.). Nach Drohungen von Facebook: AlgorithmWatch sieht sich gezwungen, Instagram-Forschungsprojekt einzustellen. *AlgorithmWatch*. <https://algorithmwatch.org/de/instagram-forschung-von-facebook-gestoppt/>
- Keller, D. (24. Februar 2022). The DSA's Industrial Model for Content Moderation. *Verfassungsblog*. <https://verfassungsblog.de/dsa-industrial-model/>
- König, P. D., & Wenzelburger, G. (2020). Opportunity for renewal or disruptive force? How artificial intelligence alters democratic politics. *Government Information Quarterly*, 37(3), 101489. <https://doi.org/10.1016/j.giq.2020.101489>
- Koster, A.-K. (2021). Das Ende des Politischen? Demokratische Politik und Künstliche Intelligenz. *Zeitschrift für Politikwissenschaft*, 32. <https://doi.org/10.1007/s41358-021-00280-5>
- Lorenz-Spreen, P., Oswald, L., Lewandowsky, S., & Hertwig, R. (2021). *Digital media and democracy: A systematic review of causal and correlational evidence worldwide*. <https://doi.org/10.31235/osf.io/p3z9v>
- Mackenzie, A. (2015). The production of prediction: What does machine learning want? *European Journal of Cultural Studies*, 18(4–5), 429–445. <https://doi.org/10.1177/1367549415577384>
- Mackenzie, A. (2017). *Machine Learners: Archaeology of a Data Practice*. The MIT Press. <https://doi.org/10.7551/mitpress/10302.001.0001>
- Makropoulos, M. (2001). *Kontingenzkultur und Optimierungsgesellschaft. Ein historisch-systematischer Aufriß*. https://www.academia.edu/36101228/Kontingenzkultur_und_Optimierungsgesellschaft_Ein_historisch_systematischer_Aufri%C3%9F
- Mol, A. (1999). Ontological politics. A word and some questions. *The Sociological Review*, 47(S1), 74–89. <https://doi.org/10.1111/j.1467-954X.1999.tb03483.x>
- Munn, L. (2022). *Automation is a myth*. Stanford University Press. https://www.sup.org/books/title/?id=34899&bottom_ref=subject
- Natale, S., & Ballatore, A. (2020). Imagining the thinking machine: Technological myths and the rise of artificial intelligence. *Convergence: The International Journal of Research into New Media Technologies*, 26(1), 3–18. <https://doi.org/10.1177/1354856517715164>
- Nowotny, H. (2021). *In AI We Trust: Power, Illusion and Control of Predictive Algorithms*. Policy Press.
- Panagia, D. (2021). On the Possibilities of a Political Theory of Algorithms. *Political Theory*, 49(1), 109–133. <https://doi.org/10.1177/0090591720959853>

- Parabel, L. (6. Februar 2020). *Künstliche Intelligenz: Niederländisches Gericht verbietet Betrugssuchmaschine*. netzpolitik.org. <https://netzpolitik.org/2020/niederlaendisches-gericht-verbietet-betrugssuchmaschine/>
- Recker, J., Berente, N., Santanam, R., & Gu, B. (2021). Managing Artificial Intelligence. *MIS Quarterly*, 45, 1433–1450. <https://doi.org/10.25300/MISQ/2021/16274>
- Rieder, B., & Hofmann, J. (2020). Towards platform observability. *Internet Policy Review*, 9(4). <https://doi.org/10.14763/2020.4.1535>
- Rieder, B., & Skop, Y. (2021). The fabrics of machine moderation: Studying the technical, normative, and organizational structure of Perspective API. *Big Data & Society*, 8. <https://doi.org/10.1177/20539517211046181>
- Rosa, H. (2020). *Gesellschaftstheorie*. UVK München.
- Rostalski, F., & Thiel, T. (2021). Künstliche Intelligenz als Herausforderung für demokratische Partizipation. Interdisziplinäre Arbeitsgruppe „Verantwortung: Maschinelles Lernen und Künstliche Intelligenz“ der Berlin-Brandenburgischen Akademie der Wissenschaften (Hrsg.), *Verantwortungsvoller Einsatz von KI? Mit menschlicher Kompetenz!* (S. 56–63). Berlin-Brandenburgische Akademie der Wissenschaften. <http://hdl.handle.net/10419/235149>
- Scheuerman, M. K., Denton, E., & Hanna, A. (2021). Do Datasets Have Politics? Disciplinary Values in Computer Vision Dataset Development. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1–37. <https://doi.org/10.1145/3476058>
- Schufa: So funktioniert Deutschlands einflussreichste Auskunftsteil (28. November 2018). *Der Spiegel*. <https://www.spiegel.de/wirtschaft/service/schufa-so-funktioniert-deutschlands-einflussreichste-auskunftsteil-a-1239214.html>
- Selbst, A. D., Boyd, D., Friedler, S., Venkatasubramanian, S., & Janet Vertesi. (2018). *Fairness and Abstraction in Sociotechnical Systems* (SSRN Scholarly Paper Nr. 3265913). Social Science Research Network. <https://papers.ssrn.com/abstract=3265913>
- Shneiderman, B. (2020). Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. *International Journal of Human-Computer Interaction*, 36(6), 495–504. <https://doi.org/10.1080/10447318.2020.1741118>
- Stark, L. (2018). Algorithmic Psychometrics and the Scalable Subject. *Social Studies of Science*, 48(2), 204–231.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.
- Thaler, R. H., Sunstein, C. R., & Balz, J. P. (2012). Choice Architecture. *The Behavioral Foundations of Public Policy* (S. 428–439). Princeton University Press. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2536504
- Ulbricht, L. (2020). Scraping the demos. Digitalization, web scraping and the democratic project. *Democratization*, 27(3), 426–442. <https://doi.org/10.1080/13510347.2020.1714595>

Valdivia, A., & de la Cueva, J. (2022). The Paradox of Efficiency: Frictions Between Law and Algorithms. *Verfassungsblog*. <https://verfassungsblog.de/roa-the-paradox-of-efficiency/>

van Deth, J. W. (2021). What Is Political Participation? *Department of Political Sociology, University of Mannheim*. <https://doi.org/10.1093/acrefore/9780190228637.013.68>