

Globale Gouvernanz Künstlicher Intelligenz und die Rolle der Schweiz

Sechs Handlungsempfehlungen

Salomé Egger

Niniane Paeffgen

Autor:innen

Salomé Egger

Salomé ist seit 2021 Direktorin des Digitalzentrum Kenia, einer von der Europäischen Union und Deutschland ko-finanzierten Initiative der Deutschen Gesellschaft für Internationale Zusammenarbeit (GIZ). Zuvor arbeitete sie als Beraterin zu Fragen rund um Digitalisierung und Menschenrechte für die deutsche Bundesregierung in Berlin und war für das United Nations Department of Economic and Social Affairs (UN DESA) in New York sowie das UNHCR in Genf tätig. Salomé hat am Graduate Institute (IHEID), SciencesPo Paris sowie an der Humboldt-Universität in Berlin studiert. Seit 2018 ist sie Autorin bei foraus in den Themenfeldern Digitalisierung und Städtediplomatie.

Niniane Paeffgen

Niniane ist Co-Gründerin des Büro für Wagemut (Philanthropy Consulting) und Mitglied des Vorstands von foraus. Zuvor war sie Geschäftsführerin der Stiftung Swiss Digital Initiative für digitale Ethik in Genf und hat in dieser Rolle die Entwicklung des ersten Labels für digitale Verantwortung vorangetrieben. Ausserdem war sie als Senior Project Manager bei digitalswitzerland für die Koordination des Beirats Digitale Transformation verantwortlich sowie für das United Nations Department of Political Affairs in New York und die internationale Handelskammer (ICC) in Paris tätig. Niniane studierte International Affairs and Governance an der Universität St. Gallen (HSG), SciencesPo Paris und schrieb ihre Masterarbeit an der Amerikanischen Universität Beirut (AUB).

Inhaltsverzeichnis

Executive Summary	2
1. Einführung	8
2. Kontext	10
2.1 Eine Technologie wie keine andere. Warum KI so mächtig ist	10
2.2 Geopolitische Implikationen: Auf dem Weg in eine technopolare Weltordnung	12
2.3 Die Suche nach einem wirkungsvollen globalen Gouvernanz-Model	14
2.4 Vorschläge zur Schaffung einer internationalen Organisation für KI-Gouvernanz	22
3. Schweizer Perspektive	24
3.1 Und was macht die Schweiz?	24
3.2 Zwischenfazit	27
4. Sechs Handlungsempfehlungen für die Schweiz	29
4.1 Kooperation im Bereich AI for Good und Responsible AI stärken	29
4.2 Das Potenzial Genfs als Standort einer möglichen neuen internationalen Organisation für KI-Gouvernanz prüfen	30
4.3 Partizipative KI-Gouvernanz und die Lokalisierung von globalen Ansätzen fördern	32
4.4 KI-Gouvernanz-Labor und Leuchtturm-Projekte für neuartige Ansätze anstossen	33
4.5 Akteure aus dem Globalen Süden stärker in internationale Debatten einbinden	34
4.6 Interner Kulturwandel und Fokussierung auf Trust & Responsible AI angehen	35
4.7 Schlusswort	36
Endnoten	37

Executive Summary

Deutsch Am 30. September 2023 erliess die Biden-Administration ein Dekret zur Regulierung von künstlicher Intelligenz (KI). Am selben Tag einigten sich die Staats- und Regierungschefs der G7 auf Leitprinzipien und einen Verhaltenskodex für die Technologie. Zwei Tage später fand der KI-Sicherheitsgipfel des Vereinigten Königreichs in Bletchley Park statt. Diese und weitere Initiativen reihen sich ein in ein zunehmend unübersichtliches Feld an KI-Gouvernanz-Ansätzen. Sie alle verfolgen das Ziel, die Entwicklung und Nutzung von KI sicherer, transparenter und ethischer zu gestalten. Denn trotz des grossen Innovationspotenzials birgt die Technologie auch ernstzunehmende Risiken für Gesellschaft und Demokratie.

Aufgrund der grenzüberschreitenden Anwendung von KI greifen nationale Regulierungsbestrebungen jedoch häufig zu kurz. Der Ruf nach internationaler Kooperation und verbindlichen Regeln für die Entwicklung und den Einsatz von KI wird daher immer lauter. Einen globalen Konsens für die Gouvernanz der Technologie gibt es bislang noch nicht – obwohl es an Vorschlägen nicht mangelt. Staatsoberhäupter, supranationale Organisationen, KI-Forschende, Tech-Unternehmen, zivilgesellschaftliche Akteure und sogar Städte haben Modelle für KI-Gouvernanz ins Leben gerufen.

Bei näherer Betrachtung stellen sich alle Ansätze aber in der einen oder anderen Weise als unzureichend heraus (geografisch, technisch, rechtlich). Nicht nur die vielfältigen Anwendungsmöglichkeiten und Eigenschaften der Technologie gestalten die Suche nach einer passenden globalen KI-Gouvernanz-Architektur als herausfordernd, sondern auch der geopolitische Wettbewerb um KI-Vorherrschaft, die zunehmend technopolare Weltordnung sowie die fehlende Inklusivität der Debatten. Zudem orientieren sich die vorgeschlagenen Ansätze oft an traditionellen Gouvernanz-Modellen, die in erster Linie auf staatliche Akteure ausgerichtet sind. KI ist aber eine Technologie, wie keine andere. Sie ist um Grössenordnungen komplexer und erfordert deshalb ein Neudenken der globalen Gouvernanz-Logiken.

Als einer der führenden Innovationsstandorte und stark vernetztes Land hat die Schweiz ein Interesse daran, sich für eine effektive globale Gouvernanz-Architektur für KI einzusetzen. Auch wenn der Handlungsspielraum begrenzt und die Schweiz bereits in relevanten internationalen Prozessen und Organisationen involviert ist, so kann sie ihre Stärken noch besser nutzen, um die globalen Leitplanken für KI mitzugestalten. Das vorliegende *Policy Paper* umreisst die aktuellen Bemühungen und Ansätze, globale Leitplanken für KI zu setzen. Es diskutiert, welche Implikationen diese für die Schweiz haben und wo Lücken bestehen. Abschliessend formulieren die Autorinnen Empfehlungen für die Schweizer Aussenpolitik und ihre Rolle in der globalen KI-Governanz. Im Fokus stehen dabei sechs strategische Handlungsfelder:

1. Kooperation im Bereich *AI for Good* und *Responsible AI* stärken
2. Das Potenzial Genfs als Standort für eine mögliche neue internationale Organisation für KI-Governanz prüfen
3. Partizipative KI-Governanz und die Lokalisierung von globalen Ansätzen fördern
4. KI-Governanz-Labor und Leuchtturm-Projekt für neuartige Ansätze anstossen
5. Akteure aus dem Globalen Süden stärker in internationale Debatten einbinden
6. Interner Kulturwandel und Fokussierung auf *Trust & Responsible AI* angehen

Français Le 30 septembre 2023, l'administration Biden a publié un décret sur la réglementation de l'intelligence artificielle (IA). Le même jour, les chef·fe·s d'État et de gouvernement du G7 se sont accordé·e·s sur des principes et un code de conduite pour cette technologie. Deux jours plus tard, c'est le sommet du Royaume-Uni sur la sécurité de l'IA qui s'est tenu à Bletchley Park. Ces initiatives ainsi que de nombreuses autres s'inscrivent dans une mosaïque de plus en plus confuse d'approches sur la gouvernance de l'IA, ayant toutes pour objectif de rendre le développement et l'utilisation de l'IA plus sûrs, plus transparents et plus éthiques. Car malgré son grand potentiel d'innovation, cette technologie présente également de sérieux risques pour la société et la démocratie.

En raison de l'utilisation transfrontalière de l'IA, les efforts nationaux de réglementation sont souvent insuffisants. L'appel à une coopération internationale et à des règles contraignantes encadrant le développement et l'utilisation de l'IA se fait donc de plus en plus pressant. Bien que les propositions ne manquent pas, il n'existe pas encore de consensus mondial pour la gouvernance de la technologie. Des chef·fe·s d'État, des organisations supranationales, des chercheur·euse·s en IA, des entreprises technologiques, des acteur·ice·s de la société civile et même des villes ont mis en place des modèles de gouvernance de l'IA.

Mais toutes les approches se révèlent insuffisantes d'une manière ou d'une autre (que ce soit sur le plan géographique, technique ou juridique). La recherche d'une architecture optimale de gouvernance mondiale de l'IA n'est pas seulement rendue difficile par la diversité des applications et des caractéristiques de la technologie, mais également par la concurrence géopolitique, l'ordre mondial de plus en plus technopolitique et le manque d'inclusivité des débats. De plus, les approches proposées s'orientent souvent vers des modèles de gouvernance traditionnels, axés en priorité sur les acteurs étatiques. Or, l'IA est une technologie comme aucune autre. Elle est d'une complexité sans pareille et nécessite donc de repenser les logiques de gouvernance globale.

En tant que centre d'innovation et pays fortement interconnecté, la Suisse a tout intérêt à s'engager en faveur d'une architecture de gouvernance mondiale efficace pour l'IA. Même si sa marge de manœuvre est limitée et qu'elle est déjà impliquée dans des processus et organisations internationaux pertinents, la Suisse peut encore davantage exploiter ses atouts afin de contribuer à la définition d'un cadre global pour l'IA. Cette publication décrit les efforts et approches actuels pour définir ce cadre mondial pour l'IA et discute de leurs implications pour la Suisse et des lacunes existantes. En conclusion, les auteures formulent des recommandations pour la politique étrangère suisse et son rôle dans la gouvernance mondiale de l'IA. Ces dernières se concentrent sur six champs d'action stratégiques :

1. Renforcer la coopération dans le domaine de *AI for Good* et *Responsible AI*
2. Examiner le potentiel de Genève en tant que siège d'une éventuelle nouvelle organisation internationale pour la gouvernance de l'IA
3. Promouvoir la gouvernance participative de l'IA et la localisation d'approches globales
4. Lancer un laboratoire de gouvernance de l'IA et un projet phare pour de nouvelles approches en la matière
5. Intégrer davantage les acteurs du Sud global dans les débats internationaux
6. Aborder le changement culturel interne et se concentrer sur le *Trust & Responsible AI*

English On 30 September 2023, the Biden administration issued an Executive Order on the regulation of artificial intelligence (AI). On the same day, the G7 leaders agreed on international guiding principles on AI and a voluntary code of conduct for AI developers. Two days later, the UK's AI Safety Summit took place at Bletchley Park. These and other initiatives compose an increasingly intricate mosaic of approaches to governing AI. They all seek to ensure that the development, deployment and use of AI advances in a way that prioritizes safety, transparency and ethical standards – because despite its vast innovative potential, the technology also carries substantial risks for societies and democracies.

Due to the widespread cross-border use of AI systems, national attempts to regulate the technology often fail. The call for international cooperation and binding rules for the development and use of AI is thus becoming ever louder. However, a unified global consensus on how to regulate the technology remains elusive – although numerous proposals have been made. Heads of state, supranational organizations, AI scholars, tech executives, civil society actors and even cities established concepts for AI governance.

Upon closer scrutiny, all approaches turn out to be insufficient in one way or another (geographically, technically, legally). The complexity of AI's multifaceted applications, coupled with its unique features, renders the search for a suitable global governance architecture extremely challenging. The unfolding geopolitical competition for AI supremacy, the emergence of an increasingly tech-driven world order, the lack of inclusivity in ongoing debates further complicate the search for global solutions. In addition, the proposed approaches are often orientated towards traditional governance models that are primarily focused on state actors. Yet AI is a technology like no other. Its complexity requires a rethinking of global governance paradigms.

As a vanguard of innovation and a nation deeply embedded in global networks, Switzerland holds a vested interest in advocating for an effective global architecture to govern AI. Even if the scope for maneuverability is limited and Switzerland is already involved in several international processes and organizations focusing on AI, the country can more strategically leverage its very own strengths

towards shaping the global framework conditions for AI. This policy paper outlines current efforts and approaches to defining global guidelines for AI. It discusses what implications these have for Switzerland and identifies gaps. The authors conclude by formulating recommendations for Swiss foreign policy and its role in global AI governance. They focus on six strategic windows of opportunity:

1. Strengthen cooperation in the field of AI for Good and Responsible AI
2. Position Geneva as host for a potential new international organization for AI governance
3. Promote participatory AI governance and the localisation of global approaches
4. Initiate an AI governance lab and lighthouse project for innovative governance ideas
5. Involve actors from the Global South closely in global debates
6. Tackle internal cultural change and focus on Trust & Responsible AI

1. Einführung

Die Einsicht, dass künstliche Intelligenz (KI) unsere Wirtschaft und Gesellschaft signifikant verändern wird, stiess in den letzten Jahren nur auf verhaltene politische Resonanz. ChatGPT kam daher einem Paukenschlag gleich, den die internationale Politik nicht überhören konnte. Die Einführung des Chatbots von OpenAI führte weltweit zu einem Wendepunkt im öffentlichen Diskurs und den politischen Bemühungen von Regierungen hinsichtlich der Entwicklung und Nutzung von KI.

Das enorme wirtschaftliche Potenzial und die vielseitigen Anwendungsmöglichkeiten von KI, im Guten wie im Schlechten, machen sie zur Schlüsseltechnologie unserer Zeit. KI ist drauf und dran, fast jede Form von Macht – sei sie wirtschaftlicher, militärischer oder politischer Natur – zu verändern. Generative KI ist dabei nur die Spitze des Eisbergs. Und damit steht die Technologie im Mittelpunkt eines geopolitischen Wettbewerbs, der den Ruf nach einer globalen KI-Gouvernanz immer lauter werden lässt. Denn KI existiert nicht in einem normativen Vakuum. Ihre (missbräuchliche) Nutzung birgt ernstzunehmende Gefahren. Insbesondere weil KI-Anwendungen oft auch grenzüberschreitend Anwendung finden und nationale Ansätze somit zu kurz greifen. Gleichzeitig entwickelt sich KI so schnell, dass Vorschriften, die heute ausgearbeitet werden, zum Zeitpunkt ihres Inkrafttretens bereits wieder veraltet sind. Künftige Auswirkungen, neue Innovationen, aber auch mögliche Risiken, können nur bedingt antizipiert werden. Je länger wir also warten, desto schwieriger wird es werden, die durch KI geschaffenen Herausforderungen anzugehen. Und je weniger Staaten bei dieser Thematik zusammenarbeiten, desto komplizierter wird die globale Durchsetzung verbindlicher Regeln und Prinzipien. Die zentrale Frage lautet also: Wie können Staaten lernen, KI erfolgreich zu steuern?

KI-Gouvernanz bezieht sich auf eine Reihe von Grundsätzen, Richtlinien und Vorschriften, welche die Entwicklung, den Einsatz und die Nutzung von KI-Systemen bestimmen. Ziel der KI-Gouvernanz ist es, sicherzustellen, dass KI auf sichere, transparente und ethische Weise entwickelt und genutzt wird. Aktuelle Regulierungsansätze zeigen, dass sie der Komplexität und dem gewaltigen Transformationspotenzial der Technologie nicht gewachsen sind. KI sprengt die Grenzen existierender Modelle globaler Gouvernanz aufgrund einer einzigartigen Kombination an Eigenschaften, die sie auszeichnet, aber auch weil bisherige Gouvernanz-Ansätze in erster Linie auf staatliche Akteure ausgerichtet sind. Bei der Entwicklung und Nutzung von KI stehen Regierungen aber fernab der eigentlichen Macht- und Entscheidungszentren, die fest in den Händen einiger wenigen Tech-Unternehmen und deren wirtschaftlichen Interessen liegen. Neuartige Gouvernanz-Ansätze werden daher dringend benötigt. Die Suche danach kommt einem Balanceakt gleich zwischen der Nutzbarmachung des riesigen Innovationspotenzials von KI bei gleichzeitiger Vermeidung möglicher negativer Auswirkungen unter Einbezug ethischer Überlegungen.

Das vorliegende *Policy Paper* umreißt die aktuellen globalen Bestrebungen und regulatorischen Ansätze, Leitplanken für KI zu setzen. Es diskutiert, welche Implikationen diese für die Schweiz haben und wo Lücken bestehen. Abschliessend formulieren die Autorinnen Handlungsempfehlungen für die Schweizer Aussenpolitik und ihre Rolle in der globalen KI-Gouvernanz. Im Fokus stehen dabei sechs strategische Windows of Opportunity.

Nota bene: Der Diskurs zur KI-Gouvernanz ist zurzeit sehr dynamisch. Das vorliegende Paper erhebt deshalb keinen Anspruch auf Vollständigkeit, sondern will einen Beitrag zu einer verbesserten Übersicht leisten und dazu beitragen, die Thematik aus einer Schweizer Perspektive besser einzuordnen.

2. Kontext

2.1 Eine Technologie wie keine andere. Warum KI so mächtig ist

Die grundlegende Prämisse von KI ist ihre Fähigkeit, kontinuierlich aus einer gesammelten Menge an Daten zu lernen. Je mehr Daten gesammelt und durch Algorithmen analysiert werden, desto besser werden die Vorhersagen der KI. Die verschiedenen Anwendungsmöglichkeiten von KI machen deutlich, dass es sich um eine sehr allgemein einsetzbare Technologie handelt, die positives Transformationspotenzial birgt, aber auch die Möglichkeit zum Missbrauch eröffnet. Das schnelle Wachstum von KI-Anwendungen kann ethische, gesellschaftliche und sicherheitsrelevante Risiken bergen. Der «Hunger» der KI nach *Big Data* steht oft im Widerspruch mit dem Recht auf Privatsphäre. KI-Modelle können auch aufgrund fehlerhafter oder nicht repräsentativer Trainingsdaten diskriminierend wirken und die Technologie kann die soziale Kohäsion und Demokratie bedrohen. So können KI-Algorithmen in den sozialen Medien bestehende Vorurteile und irreführende Inhalte verstärken und Nutzer:innen vor alternativen Standpunkten abkapseln (Stichwort: *Filter Bubbles*). Und während KI zwar bei den globalen Bemühungen um eine Verringerung der CO₂-Emissionen behilflich sein kann, verursachen grosse Sprachmodelle (*large language models*; LLM) gleichzeitig selbst signifikante Kohlenstoffemissionen durch ihren “Durst” nach noch mehr Rechenleistung.

All dies sind bereits gute Gründe, Leitplanken für KI zu fordern. KI bringt jedoch noch eine Kombination von sechs weiteren Eigenschaften mit sich, die sie von anderen Technologien unterscheidet: (1) Bei vielen ihrer Anwendungen sind die Grenzen zwischen zivilem und militärischem Nutzen fließend (*dual use*). Die gleichen Systeme, die Autos antreiben, können auch für Panzer eingesetzt werden. Dies ist deshalb so besorgniserregend, weil (2) KI eine Software ist, die vergleichsweise leicht über das Internet kopiert oder gestohlen werden kann. Man stelle sich z.B. die Proliferation einer mächtigen *Deepfake*-Anwendung vor, die von Terroristen für ihre Zwecke missbraucht wird.¹ Ein zentraler Treiber dabei ist, dass (3) die Kosten für KI-Systeme schon jetzt stark sinken, u.a. durch die Bereitstellung von KI-Modellen als Open-Source Software. Wofür führende KI-Labs heute noch Hunderte von Millionen Schweizer Franken an Entwicklungskosten investieren müssen, wird in wenigen Jahren kostenlos im Internet verfügbar sein. Gleichzeitig hat sich die Rechenleistung, die benötigt wird, um ein KI-Modell für dieselbe Aufgabe zu trainieren, seit 2012 alle 16 Monate um den Faktor zwei verringert.² Diese (4) Ubiquität von KI wiederum erhöht den Schaden, den diese anrichten kann, um ein Vielfaches. Ebenso ist die (5) Geschwindigkeit einzigartig, in der sich KI weiterentwickelt. Die fortschrittlichsten KI-Modelle von heute wenden die fünf milliardenfache Rechenleistung der besten KI-Systeme von vor zehn Jahren auf.³³ Dazu kommt schlussendlich, dass (6) KI selbstlernend ist. Wie viele unentdeckte Fähigkeiten und bisher ungeahnte Auswirkungen die Technologie entwickeln bzw. mit sich bringen wird (und bis wann), ist nicht absehbar.

KI ist nicht die erste Technologie mit einigen dieser mächtigen Eigenschaften. Sie ist aber die erste, die sie alle vereint. Jedes einzelne dieser Merkmale würde die uns bekannten Gouvernanz-Modelle in Frage stellen; alle zusammen machen sie unzureichend.

2.2 Geopolitische Implikationen: Auf dem Weg in eine technopolare Weltordnung

Das *Times Magazin* kürte kürzlich die “100 einflussreichsten Menschen” im Bereich KI. Die leitende Redakteurin beschreibt das Ergebnis als: «eine Landkarte der Beziehungen und Machtzentren, die KI-Entwicklungen vorantreiben». ⁴ Das Interessanteste an der Aufstellung? Sie zählt nur fünf Politiker:innen. Die *Times*-Liste offenbart einen grossen geopolitischen Umbruch: Unsere Weltordnung ist nicht mehr unipolar, bipolar oder multipolar, nein, wir befinden uns auf dem Weg in die Technopolarität. Technologieunternehmen verfügen inzwischen über eine solche wirtschaftliche und (sicherheits-) politische Macht, dass sie selbst zu einem geopolitischen Akteur geworden sind. Gleichzeitig ist es Regierungen bis heute nicht gelungen, diese Unternehmen und ihr Handeln effektiv zu regulieren.

Das Beispiel von Elon Musk im Kontext der Bereitstellung bzw. Abschaltung seines Sattelitendienstes *Starlink* für die ukrainische Armee zeigt eindrücklich, ⁵ wie Entscheidungen von Tech-Unternehmen nicht nur das Leben von Millionen Menschen beeinflussen, sondern auch Nationen zum Spielball ihrer Interessen machen. Sie werden zu souveränen Akteuren im digitalen Raum, den sie geschaffen haben. KI beschleunigt diesen Trend weit über die virtuelle Welt hinaus. Es ist mittlerweile für Regierungen fast unmöglich geworden, in einem angemessenen Tempo effektive Rahmenbedingungen für KI zu schaffen. Die Geschwindigkeit demokratisch legitimierter Regulierung steht immer mehr diametral zur Geschwindigkeit, in der sich KI weiterentwickelt. Die Konsequenz? Je länger wir warten, desto schwieriger wird es, die durch KI geschaffenen Probleme zu lösen. ⁶ Und je weniger Staaten in dieser Frage zusammenarbeiten, desto schwieriger wird die globale Durchsetzung verbindlicher Regeln und Prinzipien.

Eine zweite geopolitische Dimension, welche die Notwendigkeit akzentuiert, politisch die richtigen Weichen zu stellen, ist das internationale Wettrennen um die Vorherrschaft im KI-Sektor. Hierbei steht für Staaten die Nutzung von KI für Repression sowie wirtschaftliche und militärische Vorteile im Vordergrund. Die USA und China – die zwei zentralen Akteure in diesem Wettstreit – sehen dabei KI-Entwicklung vor dem Hintergrund des *Winner takes it all*-Prinzips. Auch ein Grund dafür, warum regulatorischen Eingriffen jeweils mit grosser Zurückhaltung begegnet wird. Und so heizen Washington und Peking die Beschleunigung der KI-Entwicklung zusätzlich an, z.B. durch ihr Ringen um die Dominanz in der Halbleiter-Produktion. Ein KI-Moratorium, wie von führenden KI-Expert:innen gefordert, um Risiken besser abschätzen zu lernen,⁷ wird als törichte einseitige Abrüstung eingestuft.

In der Gesamtschau sieht sich die Suche nach einer globalen KI-Gouvernanz also auch noch zwei geopolitischen Herausforderungen gegenüber, die es in einem gesamtheitlichen Rahmen zu vereinen gilt: Einerseits die Frage, wie Tech-Unternehmen reguliert werden können und andererseits wie dem derzeitigen zwischenstaatlichen Wettbewerb um KI-Vorherrschaft Einhalt geboten werden kann.

2.3 Die Suche nach einem wirkungsvollen globalen Gouvernanz-Model

Angesichts der von Nationalstaaten erwarteten erheblichen Auswirkungen von KI auf die globale Verteilung der wirtschaftlichen, militärischen und politischen Macht, versuchen derzeit Regierungen weltweit das "KI-Regulierungsvakuum" zu füllen. Denn trotz des grenzüberschreitenden Charakters der Technologie gibt es aktuell keinen einheitlichen internationalen Ansatz für deren Gouvernanz. Insgesamt handelt es sich um ein unübersichtliches Feld, welches von verschiedenen Akteuren besetzt wird. Es existiert eine Vielzahl von Initiativen (von KI-Strategien über *Whitepapers* zu Empfehlungen und *Guidelines* auf unterschiedlichsten Ebenen) von denen sich inhaltlich einige mit der Regulierung sehr spezifischer KI-Anwendungen und andere mit den allgemeinen Grundsätzen der KI-Ethik befassen. Global gesehen gibt es über 1000 Policy-Initiativen aus 60+ Ländern, wobei die meisten nach 2016 eingeführt wurden.⁸ Gleichzeitig haben auch eine Reihe an öffentlichen Institutionen, internationalen Gremien, zivilgesellschaftlichen Organisationen und Industriegruppen Gouvernanz-Ansätze mit globalem Anspruch entworfen. Es gilt zudem festzuhalten, dass Regierungen und ihre KI-Politik nicht völlig im luftleeren Raum, sondern innerhalb bestehender Rechtsstrukturen agieren und bestehende Regeln (z.B. Datenschutzgesetze) durchaus auch auf KI Anwendung finden können, leider aber oft nur begrenzt.⁹ Auf der Suche nach einem wirkungsvollen globalen Regelwerk sind die oben genannten Herausforderungen, die politischen Entscheidungsträger:innen aus China und den USA an einen Tisch zu bringen sowie Tech-Unternehmen stärker zu regulieren, nur zwei von vielen Hürden. Eine wirkungsvolle globale KI-Gouvernanz umfasst auch, dass die Bedürfnisse von Regionen berücksichtigt werden, die auf dem internationalen Parkett oft unterrepräsentiert sind, darunter Afrika, die Karibik, Lateinamerika und teilweise Asien. Gegenwärtig werden die relevanten Initiativen mehrheitlich

von Akteuren des Globalen Nordens vorangetrieben und es besteht das Risiko, dass sich ein noch grösserer Graben auftut zwischen den unterrepräsentierten Regionen und den Industrienationen. Schliesslich ist es sehr wahrscheinlich, dass der Globale Süden die Hauptlast der negativen Folgen des KI-Zeitalters tragen wird; von ausbeuterischen Klickarbeiter:innen-Jobs, die für die Herstellung von LLMs benötigt werden, bis hin zu extraktiven *Data-Mining*-Praktiken.¹⁰

Weiter stellt sich die übergeordnete Frage, was überhaupt reguliert werden soll. Da KI eine so facettenreiche Technologie ist, wird weder ein sektoraler noch eine *one-size-fits-all*-Ansatz funktionieren bzw. in einigen Fällen zu viel und in anderen zu wenig greifen. Gleichwohl muss darüber nachgedacht werden, wie eine enge Verknüpfung zwischen globaler und lokaler KI-Gouvernanz geschaffen werden kann. Auch vor dem Hintergrund, dass einige der innovativsten KI-Gouvernanz-Initiativen heutzutage von Städten entwickelt werden und die Herausforderungen und Chancen von KI sich insbesondere in Städten materialisieren. Während globale Dialoge also eine Plattform für die Erörterung übergreifender Prinzipien und Regelwerke bieten, erfordern die Dynamik und die kontextbezogenen Feinheiten von KI-Anwendungen ebenso die Miteinbeziehung lokal ausgerichteter Gouvernanz-Ansätze.¹¹

Box: Überblick relevanter KI-Gouvernanz-Initiativen

OECD, G7, G20

2019 hat die OECD eine Reihe an Prinzipien für vertrauenswürdige KI entwickelt, um einen menschenzentrierten Ansatz zu fördern. Im gleichen Jahr haben die G20 ebenfalls Grundsätze zum Umgang mit KI basierend auf denen der OECD verabschiedet. Die G7 wiederum haben im Mai 2023 den *Hiroshima AI Process* initiiert mit dem Ziel, ein einheitliches aber nicht bindendes internationales Regelwerk für die Technologie zusammen mit der OECD sowie der *Global Partnership on Artificial Intelligence* (GPAI), einer von 15 Staaten gesponserten Initiative zur verantwortungsvollen Entwicklung und Nutzung von KI, zu erarbeiten. Am 30. Oktober 2023 gaben die Staats- und Regierungschefs der G7 bekannt, dass sie sich im Rahmen des Hiroshima-KI-Prozesses auf eine Reihe von internationalen Leitprinzipien für KI und einen freiwilligen Verhaltenskodex für KI-Entwickler:innen geeinigt haben.¹²

UN-System

Der UNO-Generalsekretär António Guterres berief im Oktober dieses Jahres ein *High Level Advisory Body on AI* ein. Das Gremium soll detaillierte Empfehlungen für die globale KI-Gouvernanz entwickeln, mit einem Fokus auf bestehende Lücken.¹³ Auch andere Agenturen des UN-Systems wie etwa die *International Telecommunication Union* (ITU) mit ihrem *AI for Good Summit* haben das Thema aufgegriffen. Zudem haben alle Mitgliedsstaaten der UNESCO 2021 die „Empfehlung zur Ethik der künstlichen Intelligenz“¹⁴ angenommen. Sie stellen das erste globale normative Instrument im Bereich KI dar, das sich auf die Förderung grundlegender Prinzipien wie Transparenz und Fairness stützt. Weiter erarbeitet das B-Tech-Projekt des *Office of the High Commissioner for Human Rights* (OHCHR) an Leitlinien und Ressourcen für den Technologiesektor (beispielsweise das *Generative AI Project*) für die Umsetzung der *UN Guiding Principles on Business and Human Rights* (UNGPs).¹⁵

Europa

Auf dem europäischen Kontinent stehen zwei Entwicklungen im Fokus: Einerseits arbeitet der Europarat an einer KI-Konvention – der voraussichtlich erste, allen Staaten weltweit zur Unterschrift offenstehende, völkerrechtlich bindende Vertrag zur Förderung vertrauenswürdiger KI. Andererseits arbeitet die EU-Kommission an einem KI-Gesetz, das momentan mit den EU-Mitgliedsstaaten verhandelt wird und noch bis Ende 2023 verabschiedet werden soll (AI Act). Darin heisst es, dass KI-Anwendungen je nach Risiko, das sie für Nutzer:innen darstellen, analysiert und klassifiziert werden sollen. Die verschiedenen Risikostufen sind dann mehr oder weniger Vorschriften unterworfen. Es handelt sich dabei um den ehrgeizigsten Regulierungsversuch für KI weltweit.¹⁶

USA

Ende Oktober 2023 erliess die Biden-Administration einen *Executive Order* (EO) zur Regulierung von KI. Während die EO sich auf den Schutz und die Sicherheit von Amerikaner:innen fokussiert, wird gleichzeitig auch ein Führungsanspruch auf globaler Ebene sichtbar. Die damit befassten Behörden werden beauftragt, ihre bilateralen und multilateralen Engagements zu intensivieren, für «robuste internationale Rahmenbedingungen» zu sorgen und die Entwicklung gemeinsamer Standards mit Partnerländern zu beschleunigen.¹⁷ Die EO ergänzt die bisherigen Bemühungen, die vor allem auf Selbstregulierung setzen. So gab die US-Regierung im Juli 2023 bekannt, dass sie von acht führenden Tech-Unternehmen freiwillige Verpflichtungen für den verantwortungsvollen Umgang mit KI erhalten habe.¹⁸

UK

Dem Vereinigten Königreich ist im November 2023 mit dem *AI Safety Summit* in Bletchley Park ein diplomatischer Coup gelungen. Bei der Zusammenkunft von globalen Tech-Expert:innen aus Industrie, Forschung und Zivilgesellschaft mit politischen Entscheidungsträger:innen wurde eine internationale Erklärung unterzeichnet, in der die Notwendigkeit anerkannt wird, sich mit den von der KI-Entwicklung ausgehenden Risiken zu befassen und etwa neuartige KI-Modelle zu testen, bevor sie zur öffentlichen Nutzung freigegeben werden. Dafür soll in der UK ein globales *AI Safety Institute* (AISi) entstehen, das die Sicherheit neuer KI-Systeme vor und nach ihrer öffentlichen Freigabe analysieren soll. Die Erklärung wird von mehr als 25 Ländern, unter anderem China, Australien und Kanada sowie der EU unterstützt.¹⁹

China

China hat in den letzten Monaten eine Reihe von Massnahmen zur Regulierung von KI ergriffen, darunter den Entwurf von Regeln für ein ethisches Prüfverfahren für die Entwicklung bestimmter KI-Modelle. Am dritten *Belt and Road Forum for International Cooperation* im Oktober 2023 kündigte Präsident Xi Jinping die *Global AI Governance Initiative* (GAIGI) für die Teilnehmerländer der Initiative an.²⁰ GAIGI fordert einen menschenzentrierten Ansatz, mehr Inklusivität sowie Fairness bei der Entwicklung von KI. Die Initiative soll alle 155 Länder umfassen, die an der *Belt and Road Initiative* beteiligt sind, und würde damit zu einem der umfassendsten KI-Gouvernanz-Foren avancieren.

ASEAN

Im Gegensatz zum EU KI-Gesetz verfolgen die südostasiatischen Länder bei der KI-Gouvernanz einen selbstregulatorischen Ansatz. So entwickelt ASEAN einen freiwilligen KI-Ethikleitfaden für Unternehmen, der kulturelle Unterschiede betont und keine strengen Risikokategorien vorschreibt. Er ermutigt Regierungen, den Privatsektor bei der ethischen Anwendung von KI zu unterstützen, überlässt aber die Einzelheiten den lokalen Regulierungsbehörden und Unternehmen.²¹

Afrikanische Union

Das *High-Level Panel on Emerging Technologies* der Afrikanischen Union (AU) und die AU Entwicklungsagentur (AUDA-NEPAD) haben im März afrikanische Expert:innen für KI zu einem Treffen in Kigali, Ruanda, zusammengerufen, um die *African Union AI Continental Strategy for Africa* fertigzustellen. Die Strategie soll den AU-Mitgliedsstaaten zur Überprüfung und Validierung vorgelegt und im Januar 2024 verabschiedet werden.

Internationale Normungsorganisationen

Weitere relevante Referenzpunkte sind internationale Standardsetzungsorganisationen (SSO) wie etwa die *International Organization for Standardization* (ISO) oder die *IEEE Standards Association*, deren Mitgliederbasis in der Regel von Industrie- und Wirtschaftsverbänden dominiert werden. Durch ihre Entwicklung von KI-Standards und ihrer institutionellen Fähigkeit, einen globalen Expertenkonsens zu erzielen, üben sie eine normsetzende Kraft aus.

Privatsektor

Eine Allianz amerikanischer Tech-Unternehmen hat die *Partnership on AI* (PAI) ins Leben gerufen, deren Mitgliedschaft kürzlich um eine Reihe von Nichtregierungsorganisationen erweitert wurde. PAI zielt darauf ab, die Forschung im Zusammenhang mit KI-Ethik und Gouvernanz zu fördern. Grosse Unternehmen wie Meta, Google, Microsoft etc. haben sich ausserdem selbst Leitlinien und Prinzipien zum Umgang mit KI gegeben oder Ethikboards dafür eingesetzt. Die Wirkung dieser selbstregulatorischen Massnahmen wird jedoch als wenig effektiv kritisiert und als Versuch gewertet, staatliche Regulierungsmassnahmen zu umgehen.

Städte

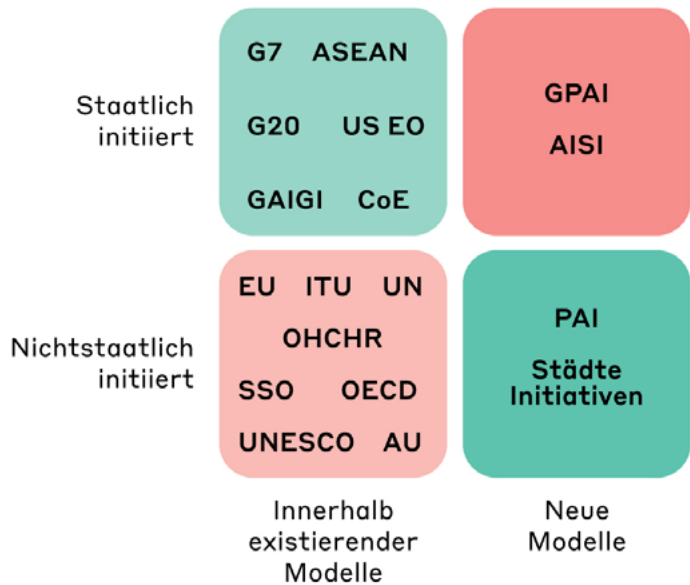
Zu den prominentesten Regulierungsinitiativen gehören die *Cities Coalition for Digital Rights*, eine 50+ Städte-Initiative mit ihrem *Global Observatory of Urban AI* oder die *Montreal Declaration for Responsible AI*, sowie andere Bemühungen, die beispielhaft dafür sind, wie Städte in der KI-Gouvernanz-Landschaft versuchen eine Vorreiterrolle zu spielen.

Es konkurrieren also zahlreiche unterschiedliche Initiativen um Einfluss in einem Flickenteppich der KI-Gouvernanz. Viele Ansätze sind staatlich geführt, aber auch internationale Organisationen und Städte haben ein signifikantes Niveau an Handlungsfähigkeit an den Tag gelegt. Analysiert man die Initiativen inhaltlich, stellt man fest, dass die meisten Ansätze das gemeinsame Engagement für vertrauenswürdige KI betonen. Es wird jedoch nicht klar, wie eine Zusammenarbeit aussehen könnte, die über die bestehenden nationalen Massnahmen und Partnerschaften mit internationalen Organisationen wie der OECD hinausreicht. In einer Welt, in der sich in Zukunft schädliche KI-Anwendungen rasant schnell ausbreiten könnten, kommt dies jedoch einem signifikanten Risiko gleich. Denn Werte und ethische Prinzipien müssen auch in ein System von Regeln und Standards, verbindlichen Vereinbarungen und internationalen Mechanismen überführt werden, die über Selbstverpflichtungen oder Empfehlungen hinausgehen.

Ordnet man die wichtigsten Initiativen den Dimensionen “staatlich vs. nichtstaatlich” sowie “aufbauend auf bereits existierenden Gouvernanz-Modellen vs. Schaffung neuer Gouvernanz-Instrumente” zu, dann wird in erster Linie auf bestehende Architekturen zurückgegriffen.

Aufgrund der beschriebenen geopolitischen Implikationen und den einzigartigen Eigenschaften von KI, braucht es aber einen neuartigen Ansatz. Denn KI zwingt uns, bestehende Annahmen über die geopolitische Weltordnung neu zu denken, jenseits des nationalstaatlichen Paradigmas. Eine erfolgreiche globale KI-Gouvernanz scheint daher zunehmend davon abhängig zu sein, ob es das internationale System schafft, die traditionelle Vorstellung von Souveränität hinter sich zu lassen. Das heisst, dass es nicht nur Tech-Unternehmen mit an den Tisch holt, sondern auch – in den Worten der Harvard-Informatikerin Barbara Grosz – die «fehlende Dimension der KI-Gouvernanz»,²² nämlich die Zivilgesellschaft und Endnutzer:innen. Denn obwohl deren Beteiligung regelmässig als Schlüssel zur Gewährleistung ethischer und fairer Ansätze für die Steuerung von KI durch staatliche und nichtstaatliche Akteure genannt wird, steht

die Realität der KI-Gouvernanz im krassen Gegensatz dazu. Und dies trotz der Erkenntnis, dass viele der Probleme, die KI-Systeme mit sich bringen, auf die mangelnde demokratische Beteiligung und die fehlende Partizipation der betroffenen Gemeinschaften, Bürger:innen, der Zivilgesellschaft und der Gesellschaft insgesamt zurückzuführen sind.²³



2.4 Vorschläge zur Schaffung einer internationalen Organisation für KI-Gouvernanz

In einem Blogbeitrag im Mai dieses Jahres vertraten führende Mitarbeitende von OpenAI die Ansicht, dass eine Art Internationale Atomenergie-Organisation (IAEO) für KI erforderlich sei, um komplexe KI-Systeme sicher zu verwalten.²⁴ In eine ähnliche Richtung geht der Vorschlag des UN-Delegierten für Digitalisierung Amandeep Gilles, der für eine eigene UN-Agentur für KI weibelt. In einem Artikel,²⁵ der im Juli von Google DeepMind, OpenAI sowie verschiedener Forschungseinrichtungen veröffentlicht wurde, werden sogar vier globale KI-Behörden vorgestellt. Neben einer IAEO-ähnlichen Agentur, die branchenweite Standards entwickeln und Akteure überwachen soll, wird eine Institution nach dem Vorbild des *Intergovernmental Panel on Climate Change* (IPCC) vorgeschlagen, die einen Expert:innenkonsens zu technischen Fragen ermöglicht. Ein Vorschlag, den EU-Kommissionspräsidentin Ursula von der Leyen im September ebenfalls aufgriff.²⁶ Zudem wird eine globale Partnerschaft zwischen dem öffentlichen und dem privaten Sektor vorgeschlagen, die den gleichberechtigten Zugang zu gemeinnütziger KI sicherstellen soll, ähnlich wie die GAVI Impfallianz. Die vierte Idee zielt auf internationale Zusammenarbeit bei der KI-Sicherheitsforschung ab, analog zum Vorgehen bei der Europäischen Organisation für Kernforschung (CERN).²⁷

Derweil schlug das Massachusetts Institute of Technology (MIT) im Juli die Schaffung eines *Global AI Observatory* (GAIO) vor mit dem Ziel, eine globale, standardisierte Datenbank zur Meldung von KI-Vorfällen aufzusetzen: Ein Register wichtiger KI-Systeme mit Schwerpunkt auf KI-Anwendungen, welche die grössten sozialen und wirtschaftlichen Auswirkungen bergen. Die GAIO soll zudem die Zusammenführung des weltweiten Wissens über die Auswirkungen von KI auf kritische Bereiche koordinieren und die globale KI-Debatte durch einen jährlichen Bericht über

den Stand der KI strukturieren.²⁸ Zuletzt veröffentlichte auch Microsoft im August einen Vorschlag zusammen mit Forschenden verschiedener akademischer und gemeinnütziger Institutionen. Es könnte eine internationale KI-Organisation gegründet werden, die mit nationalen Regulierungsbehörden für die Festlegung von Standards zusammenarbeitet und Länder zertifiziert, anstatt einzelne KI-Unternehmen zu überwachen. Im Rahmen dieses Modells würden die einzelnen Regierungen die Standards durchsetzen, indem sie nur Unternehmen aus Staaten, die diese Standards erfüllen, den Zutritt zu ihrem Markt gewähren.²⁹

Während Mitgliedschaft und Mandat unterschiedlich ausfallen, scheint sich doch über das Jahr 2023 hinweg eine Tendenz in Richtung Schaffung einer oder mehrerer komplementär agierender internationalen Organisation(en) für KI-Gouvernanz herauszukristallisieren. Die Mehrheit der Ideen wird hierbei von privatwirtschaftlichen und/oder akademischen Akteuren propagiert. Es wird sich zeigen, ob und wie Staaten in ihren Bemühungen um globale KI-Gouvernanz diese aufgreifen werden.

3. Schweizer Perspektive

3.1 Und was macht die Schweiz?

In einem 2019 veröffentlichten Bericht zu den “Herausforderungen der Künstlichen Intelligenz” kam eine interdepartementale Arbeitsgruppe des Bundes zum Schluss, dass der bestehende generelle Rechtsrahmen in der Schweiz derzeit grundsätzlich geeignet ist, mit neuen Anwendungen und Geschäftsmodellen im Bereich KI umzugehen.³⁰ Ausserdem erliess der Bundesrat, neben den Strategien “Digitale Schweiz” und “Digitalaussenpolitik”³¹ Leitlinien für den Umgang mit KI in der Bundesverwaltung.³² Die Leitlinien sind als Orientierungsrahmen für den Umgang mit der Technologie im Bund gedacht und gemäss einem 2022 durchgeführten Monitoring besteht derzeit kein Bedarf, die Leitlinien anzupassen oder zu aktualisieren.³³

Um den Austausch über KI zu gewährleisten, existieren innerhalb der Bundesverwaltung zudem verschiedene Gremien. Seit 2019 wird die vom Bundesamt für Kommunikation (BAKOM) koordinierte *Plateforme Tripartite* genutzt, um politische, gesellschaftliche und wirtschaftliche Fragen rund um KI mit über 300 Expert:innen aus Verwaltung (Bund und Kantone), Wirtschaft, Zivilgesellschaft und Wissenschaft zu diskutieren. Weiter schuf der Bundesrat 2021 ein Kompetenznetzwerk KI (CNAI) für technische Aspekte der konkreten Nutzung von KI in der Bundesverwaltung und prüft die Gründung eines “Knotenpunkts Recht” für alle rechtlichen Aspekte. Der “Knotenpunkt Algorithmen” unter der Leitung des *Data Science Competence Centre* (DSCC) des Bundesamts für Statistik stellt dem öffentlichen Sektor seine Expertise in der konkreten Anwendung von Datenwissenschaft und KI zur Verfügung. Weiter verfolgt und analysiert eine interdepartementale Koordinationsgruppe unter der Leitung des BAKOM und des Staatssekretariats des Eidgenössisches Departement für auswärtige Angelegenheiten (EDA)³⁴ die EU-Digitalpolitik und Arbeiten am KI-Gesetz der EU.

Grundsätzlich fährt die Schweiz eine Strategie der liberalen und technologieutralen³⁵ Gesetzgebung und sieht die Industrie und Wissenschaft primär in der Pflicht, ihre Verantwortung wahrzunehmen. Aus diesem Grund hat man sich gegen eine nationale KI-Strategie entschieden. Es wird eine beobachtende und abwartende Haltung eingenommen. Die Bundesverwaltung wird bis Ende 2024 verschiedene Ansätze für die Regulierung von KI evaluieren und anschliessend über nächste Massnahmen entscheiden. Es ist allerdings zu erwarten, dass sich die Schweiz an europäischen Standards orientieren wird, insbesondere am KI-Gesetz der EU. Der eidgenössische Datenschutz- und Öffentlichkeitsbeauftragte (EDÖB) bezog ausserdem Stellung, dass unabhängig von möglichen Regulierungsmassnahmen, das aktuelle Datenschutzgesetz technologieneutral formuliert und demzufolge auch auf den Einsatz von KI-gestützten Datenbearbeitungen direkt anwendbar ist.³⁶ Zusätzlich hat der Bundesrat beim Eidgenössischen Departement für Umwelt, Verkehr, Energie und Kommunikation (UVEK), unter Einbezug aller anderen Departemente, eine Übersicht möglicher Regulierungsansätze von KI in Auftrag gegeben. Damit soll die Basis für einen konkreten Auftrag für eine Regulierungsvorlage zu KI im Jahr 2025 geschaffen werden.³⁷

Soweit der *Status Quo* auf nationaler Ebene. Mit Blick auf die internationale Ebene, hat die Schweiz als kleines, hochentwickeltes und vernetztes Land ein Interesse daran, die Debatte über die globale Gouvernanz von KI mitzugestalten. Sie hat sich bislang aktiv in die Debatten um Leitplanken für KI eingebracht u.a. durch die Mitwirkung in einschlägigen internationalen Organisationen und Prozessen wie dem Europarat, wo sie aktuell den Vorsitz des KI-Komitees innehat und an den Verhandlungen für ein erstes verbindliches internationales Abkommen zu KI beteiligt ist. Im Falle einer Ratifikation des verhandelten Abkommens durch sie wäre die Schweiz danach auch zur innerstaatlichen Umsetzung dieses Abkommens verpflichtet. Die Schweiz war auch Mitglied der OECD-Expertengruppe, welche die 2019 angenommenen OECD-Grundsätze für KI verfasst hat. Der durch die ITU organisierte *AI for Good Summit* in Genf, an dem

zentrale Gouvernanz-Ansätze diskutiert und relevante Akteure (auch aus der Privatwirtschaft) zusammengebracht werden, hat zudem internationale Ausstrahlung weit über Genf hinaus.

Weiter hat die Schweiz im Juli 2023 im UN-Sicherheitsrat auf die Notwendigkeit eines gemeinsamen Rahmens für alle Beteiligten bei der Entwicklung und Anwendung von KI verwiesen und die Erfordernisse von mehr Transparenz in Bezug auf ihre Funktionsweise und die verwendeten Daten gefordert. In der *Swiss Call for Trust & Transparency Initiative* suchen Wissenschaft, Privatsektor und Diplomatie ausserdem gemeinsam nach praktischen und schnellen Lösungen für KI-basierte Risiken.³⁸ Und auch 2019 gab es bereits Bestrebungen, Vertrauen und Ethik in neue Technologien zu stärken: Als konkretes Ergebnis aus dem *Swiss Global Digital Summit for Ethics and Fairness in the Digital Age* wurde die *Swiss Digital Initiative* gegründet mit dem Ziel, ethische Prinzipien in der Digitalisierung zu operationalisieren.³⁹

Aufgrund ihrer Neutralität, ihres Rufes als Forschungs- und Innovationsstandorts sowie der Bedeutung Genfs als Standort für globale Gouvernanz, hat die Schweiz signifikantes Potenzial hinsichtlich der Mitgestaltung und Koordinierung von Fragen der KI-Governanz. Im internationalen Genf sind zahlreiche Organisationen und Akteure ansässig, die als *Centre of Normative Power* gelten. Das sind Akteure wie UN-Mitgliedsstaaten, multilaterale Organisationen oder private Unternehmen, die eine wichtige Rolle in der Gestaltung für weltweit geltende Normen und Standards bei KI-Systemen spielen. Allerdings werden die *Windows of Opportunity* für die Schweiz, international eine gestaltende Rolle zu spielen, zusehends kleiner – zu KI im Spezifischen und in Digitalisierungsfragen ganz allgemein. Etliche Institutionen der in Genf ansässigen internationalen Organisationen haben ihre digitalen Ableger bereits woanders aufgebaut. Etwa das Internationale Komitee vom Roten Kreuz (IKRK) in Luxemburg, das *United Nations International Computing Centre* (UNICC), der technische Dienstleister der UNO, der ein grosses Datenzentrum in Valencia aufbaut, oder die Weltgesundheitsorganisation (WHO), die einen *Hub for Pandemic and Epidemic Intelligence* in Berlin entwickelt.⁴⁰

3.2 Zwischenfazit

Es wird zunehmend schwieriger, den Überblick über die zahlreichen Initiativen und Vorstösse für eine globale KI-Gouvernanz zu behalten. KI entwickelt sich in rasender Geschwindigkeit und Vorschriften, die heute ausgearbeitet werden, werden zum Umsetzungszeitpunkt bereits wieder veraltet sein und zukünftige Schäden und Innovationen möglicherweise nicht abdecken.

Ferner weisen die bisherigen Ansätze Gouvernanz-Lücken auf: (a) Viele sind geografisch unzulänglich, d.h., es handelt sich um regionale Bemühungen, die aufgrund ihrer Mitgliedschafts- und Beteiligungspolitik derzeit nicht offen und inklusiv genug sind, um als wirklich globale Plattform für die benötigte Koordination zu dienen. Besonders die Akteure des Globalen Südens bleiben aussen vor. Die Bemühungen der UNESCO sind angesichts der hohen Beteiligung von KI-Expert:innen weltweit und den zwischenstaatlichen Verhandlungen der bisher globalste Ansatz. (b) Andere Modelle sind inhaltlich unscharf, da sie sich entweder nur auf einen spezifischen Bereich von KI-Anwendungen fokussieren oder aber so allgemein gehalten sind, dass konkrete technische Fragestellungen unbeantwortet bleiben. (c) Die allermeisten Ansätze sind zudem rechtlich zahnlos, weil die Durchsetzungsmechanismen fehlen und beispielsweise nur als Prinzipien für vertrauenswürdige KI formuliert sind. Eine Übersetzung in konkrete Umsetzungsmassnahmen gibt es nicht. (d) Ausserdem orientiert sich eine Grosszahl der Initiativen zu stark an bisherige Gouvernanz-Vorstellungen. (e) Zu guter Letzt holen die wenigsten der vorgeschlagenen Modelle neben Tech-Unternehmen auch proaktiv die Zivilgesellschaft und Endnutzer:innen mit einem ernsthaften Mitspracherecht an den Verhandlungstisch.

Was bedeutet dies im Umkehrschluss? Damit Staaten KI erfolgreich steuern können, muss eine KI-Gouvernanz-Architektur geschaffen werden, die nicht nur von einem globalen Konsens getragen wird (oder als Minimalstandard von den USA und China), sondern auch technische Standards mitdenkt und gleichzeitig auf möglichst viele KI-Anwendungen übertragbar ist. Zugleich muss sie Prinzipien in rechtlich verbindliche Regelungen überführen, agil genug sein, um den einzigartigen Charakteristika von KI zu begegnen und auf einem Multi-Stakeholder-Ansatz beruhen, der zivilgesellschaftliche Akteure sowie Tech-Unternehmen sinnvoll mit ins Boot holt. Dass die Schaffung eines solchen Regelwerks in einem zunehmend polarisierten politischen Klima einem sehr ambitionierten Unterfangen gleichkommt, ist dabei ein *Understatement*.

Dennoch sollte sie das übergeordnete Endziel sein. Wenngleich vorübergehend davon auszugehen ist, dass es in naher Zukunft nicht *die* eine Lösung oder *das* eine Instrument für den Einsatz von KI geben wird, sondern einen "KI -Gouvernanz-Baukasten". Also ein Set an Standards und Prinzipien sowie Vorschriften auf verschiedenen technischen Steuerungsebenen von KI in Kombination mit mehr oder weniger konkreten Umsetzungsmechanismen und rechtlichen Durchsetzungsmöglichkeiten, die sich jedoch regional unterscheiden dürften. Insbesondere hinsichtlich rechtlicher Handhabung werden sich die meisten Länder grösstenteils an Europa orientieren und abwarten, wie sich das neue KI-Gesetze in der Praxis auswirkt. Es ist ebenfalls davon auszugehen, dass es verschiedene Arten von KI-Regulierungsbehörden geben wird, die den Auftrag innehaben werden, die existierenden Dilemmata zu diskutieren und der Öffentlichkeit zu erklären. Weiter bleibt es spannend zu beobachten, ob sich in den nächsten Jahren eine internationale Organisation für KI-Gouvernanz etablieren und wie breit ihr Mandat sein wird.

Das grösste Risiko mit Blick auf das übergeordnete Ziel einer globalen KI-Gouvernanz-Architektur besteht zurzeit insbesondere darin, dass die Vielzahl an Bemühungen unverbindlich und fragmentiert bleiben und keine Lösung ausreichend Zugkraft entwickelt wird, um konsolidierend in eine Richtung zu wirken und die identifizierten Lücken zu füllen.

4. Sechs Handlungsempfehlungen für die Schweiz

Der Handlungsspielraum für die Schweiz im Bereich KI-Gouvernanz ist begrenzt. Dennoch gibt es einige *Windows of Opportunity*, welche sie strategisch nutzen könnte. Die nachfolgenden sechs Handlungsempfehlungen zeigen auf, wie die Schweiz ihr Engagement im Bereich der globalen KI-Gouvernanz weiter stärken und sich für verantwortungsvolle KI-Anwendungen einsetzen kann:

4.1 Kooperation im Bereich AI for Good und Responsible AI stärken

Ein Weg, um stärkere Konvergenz in den globalen KI-Gouvernanz-Bemühungen herbeizuführen, ist, sich stärker dafür einzusetzen, Bereiche für kooperative Projekte in der KI-Forschung und auf der Anwendungsebene zu finden. Das wohl grösste Potenzial für Kooperation liegt dabei in gemeinwohlorientierten Initiativen, die sich mit globalen Problemen wie dem Klimawandel oder dem Erreichen der Ziele für nachhaltige Entwicklung (*Sustainable Development Goals*, SDGs) beschäftigen. Mit dem *Arc Lémanique* als *Melting Pot* für zentrale Stakeholder in den Bereichen Nachhaltigkeit, Digitalisierung und Gouvernanz (bspw. UNO, ITU, DiploFoundation, CyberPeace Institute, GESDA, Swiss Digital Initiative, EPFL etc.) sowie den in Zürich ansässigen Tech-Firmen und der ETH kann die Schweiz auf ein global agierendes digitales Ökosystem zurückgreifen

und sollte dies strategisch dafür nutzen, um eine internationale Drehscheibe für *AI for Good*-Anwendungen aufzubauen. So könnte die Schweiz z.B. weitere Kohorten der *AI for Good Innovation Factory*⁴¹ der ITU unterstützen, über die Start-ups und Unternehmer:innen weltweit KI-Innovationen für die SDGs pitchten. Sie kann auch globale Wissenschaftskooperationen zu KI und den SDGs über ihr Swissex-Netzwerk etablieren. Und mit dem *AI for Good Summit* der ITU in Genf oder dem Weltwirtschaftsforum (*World Economic Forum, WEF*) in Davos besitzt die Schweiz relevante Plattformen, um die neuesten Erkenntnisse und gemeinnützigen KI-Anwendungen öffentlichkeitswirksam vorzustellen und mit relevanten Stakeholdern zu diskutieren. Die Schweiz könnte ausserdem zu einem Multistakeholder-Mapping-Prozess analog zum Montreux-Prozess einladen, wie dies in einem vorgängigen foraus-Diskussionspapier vorgeschlagen wurde.⁴²

4.2 Das Potenzial Genfs als Standort einer möglichen neuen internationalen Organisation für KI-Gouvernanz prüfen

Die Schweiz sollte sich auf dem internationalen Parkett von Anfang an in die Diskussionen rund um die Schaffung einer neuen internationalen Organisation für KI-Gouvernanz einbringen und die Möglichkeit prüfen, gegebenenfalls als Gaststaat einer solchen Organisation zu fungieren. Dies würde ebenfalls auf die Ambitionen Genfs einzahlen, als internationale Hauptstadt der Digitalgouvernanz zu gelten und weiter die Rolle als Gaststaat internationaler Organisationen und als wichtiges Zentrum für globale Gouvernanz festigen.⁴³ Als neutralen Standort mit einem exzellenten Innovationsökosystem könnte die Schweiz damit einen relevanten daten- und faktenbasierten Beitrag zur globalen Debatte leisten. Dies nicht nur in traditionellen Bereichen wie Frieden und der humanitären Arbeit, sondern auch wenn es um Zukunftstechnologien geht.

Wenn man sich die bislang vorgeschlagenen möglichen Mandate einer solchen Organisation ansieht (siehe Kapitel 4), dann bringt Genf eine Reihe an Grundvoraussetzungen mit, um erfolgreich eine mögliche globale KI-Institution zu beherbergen. Sollte es auf eine IAEO⁴⁴ – oder IPPC⁴⁵ – ähnliche Agentur hinauslaufen, mit dem Mandat branchenweite Standards zu entwickeln und Expert:innenkonsens zu technischen Fragen herbeiführen, dann können ITU und ISO als zentrale Normungsorganisationen als Sprungbrett dienen, da sie im internationalen Regelwerk für Digitalisierung eine Scharnierfunktion innehaben.

Sollte sich eher eine Behörde im Sinne der GAIO⁴⁶ nach dem Vorschlag des MIT herauskristallisieren, dann wären EPFL, ETH und die Universität Genf ein erfolgversprechendes Trio, um eine standardisierte Datenbank zur Meldung von KI-Vorfällen aufzusetzen sowie ein Register zentraler KI-Anwendungen mit den signifikantesten sozioökonomischen Auswirkungen zu erstellen. Gleichzeitig hat Genf mit der DiploFoundation bereits heute einen kompetenten Aggregator bezüglich globalem Wissen zu den Effekten von KI auf kritische Bereiche, der auch glaubwürdig den vorgeschlagenen jährlichen “GAIO-Bericht” über den Stand der KI veröffentlichen und damit die globale KI-Debatte inhaltlich strukturieren könnte. Ein möglicher Anknüpfungspunkt wäre ausserdem die Zusammenarbeit mit dem neu in den UK anzusiedelnden *AI Safety Institute*.

Eine Machbarkeitsstudie wäre ein erster Schritt, um die Vorzüge Genfs als Gaststaat für eine neue internationale Organisation für KI-Gouvernanz *en détail* herauszuarbeiten. Dies insbesondere vor dem Hintergrund, dass der Bundesbeschluss über die „Massnahmen zur Stärkung der Rolle der Schweiz als Gaststaat 2020–2023“⁴⁷ ausläuft und neu verhandelt werden dürfte. Ein stärkerer Fokus in den neu zu definierenden Massnahmen auf neue Technologien und konkret auf die Gouvernanz von KI sind zentral, wenn Genf und die Schweiz in diesen Fragen zukünftig noch eine relevante Rolle spielen möchten.

4.3 Partizipative KI-Gouvernanz und die Lokalisierung von globalen Ansätzen fördern

Während nationale und supranationale Gouvernanz-Strukturen mit der Etablierung von KI-Rahmenwerken ringen, erlassen Städte und urbane Regionen bereits Gesetze und Leitlinien für den Umgang mit der Technologie. Die Schweiz könnte sich durch Städte wie Genf und Zürich darum bemühen, formelle oder informelle Mechanismen zu schaffen, welche diesen *AI Localism* mit den internationalen Anstrengungen für KI-Gouvernanz verknüpfen und somit auch hier zu einer stärkeren Konvergenz der Diskussionen beitragen. Städte, wie Genf, Zürich, Bern etc., könnten auch den urbanen Dialog um die Frage vorantreiben, wie Städtediplomatie und das lokale Fachwissen als wirksame *Bottom-up*-Instrumente eingesetzt werden können, um globale Regelungen zu ergänzen bzw. zu lokalisieren. Hierbei können die Städte bspw. mit der durch das WEF angeleitete *G20 Global Smart Cities Alliance on Technology Governance* zusammenarbeiten. Ferner könnten Genf und Zürich der *Cities Coalition for Digital Rights* beitreten, einem Städtenetzwerk, das darauf ausgerichtet ist, digitale Herausforderungen zu lösen und rechtliche, ethische und operative Rahmenbedingungen zu erarbeiten, die Menschenrechte im digitalen Raum stärken. Hierbei könnten sie sich insbesondere darauf fokussieren, das *Global Observatory of Urban AI* des Netzwerks zu unterstützen und im internationalen Genf besser zu positionieren.

Ergänzend dazu kann sich die Schweiz (bspw. mithilfe von Städten wie Genf und Zürich) für eine partizipative KI-Gouvernanz einsetzen, indem sie die Sichtweisen und Bedarfe von Bürger:innen und Endnutzer:innen sammelt, aggregiert und in die Debatten auf internationaler Ebene einfließen lässt. Dies könnte durch das Instrument der *Citizen Assembly* geschehen, das bereits andernorts erfolgreich eingesetzt wurde.⁴⁸ Bürgerversammlungen bringen nach dem Zufallsprinzip ausgewählte Bürger:innen an einem (virtuellen) Ort von einem Tag bis zu mehreren Wochen

zusammen. Versammlungsmitglieder hören von Expert:innen über ein spezifisches Thema und beraten sich danach ausführlich, bevor sie politische Empfehlungen oder Vorschläge formulieren. Ihr Reiz liegt insbesondere darin, dass sie ein einzigartiges Spektrum an Stimmen und Lebenserfahrungen einfangen und so politische Empfehlungen hervorbringen, die den Präferenzen der breiten Bevölkerung besser entsprechen, mit grösserer Wahrscheinlichkeit akzeptiert werden und Vertrauen in KI schaffen. Die Schweiz könnte eine oder mehrere internationale KI-Bürgerversammlung(en) organisieren und dabei Menschen aus dem Globalen Norden und Süden zusammenbringen. Inhaltlich würde sich ein Gouvernanz-Fokus auf KI-Anwendungen anbieten, welche die grössten Menschenrechtsrisiken bergen wie z.B. der automatisierte Zugang zu staatlichen Dienstleistungen oder die Erfassung, Speicherung und Verarbeitung von biometrischen, gesundheitlichen oder anderen sensiblen persönlichen Daten. Dank jahrzehntelanger Erfahrung mit Öffentlichkeitsbeteiligung in anderen Themenfeldern, gibt es bereits eine Reihe an *Best Practices* und *Lessons Learnt*, auf die zurückgegriffen werden kann oder welche bereits speziell für KI-Gouvernanz aufbereitet wurden (siehe bspw. Praxisleitfaden für aussagekräftige öffentliche Beteiligung zu KI von Michele Gilman, University of Baltimore).⁴⁹

4.4 KI-Gouvernanz-Labor und Leuchtturm-Projekte für neuartige Ansätze anstossen

Die Schweiz könnte sich weiter für die Entwicklung von innovativen Gouvernanz-Modellen einsetzen, d.h., zusammen mit Akteuren aus Wirtschaft, Gesellschaft und Forschung ko-kreativ neuartige, über die in Kapitel 4 hinausgehende Lösungsansätze sammeln und weiterentwickeln und sich insbesondere auf bestehende Lücken fokussieren. Vor- und Nachteile der jeweiligen Gouvernanz-Ideen könnten diskutiert, strukturiert und in Form einer *Global AI Governance*

Idea Map öffentlich zugänglich gemacht werden, um (politischen) Entscheidungsträger:innen Denkanstösse zu geben, die über die vertrauten Gouvernanz-Ansätze hinausgehen. Das Ideenlabor könnte auch Zukunftsszenarien basierend auf Methoden der strategischen Vorausschau für die neuen Gouvernanz-Modelle entwerfen, welche dabei helfen, diese besser zu verstehen und greifbarer zu machen. Ein solches Ideenlabor kann beispielsweise einem Zusammenschluss von technischen Hochschulen (Tech) und dem Graduate Institute (Politik) entspringen, durch *Design Thinking*-Spezialist:innen beraten und in internationale Prozesse durch das WEF eingebunden werden. Wenn es der Schweiz ausserdem gelingen würde, ein Leuchtturm-Projekt aufzubauen, das eine der sechs diskutierten Gouvernanz-Lücke adressiert, so könnte damit ein Mehrwert erzielt werden, der über Konferenzen oder Forschungsaktivitäten hinausgeht.

4.5 Akteure aus dem Globalen Süden stärker in internationale Debatten einbinden

Die Länder des Globalen Südens werden voraussichtlich am wenigsten von KI profitieren bzw. die Hauptlast der negativen Folgen tragen. Viele Entwicklungs- und Schwellenländer sind zudem mit einem Spektrum von KI-bezogenen Herausforderungen konfrontiert, die in den Industrieländern weniger häufig anzutreffen sind. Die Einbeziehung von Expert:innen aus diesen Ländern kann daher wertvolle zusätzliche Einblicke in die vielschichtigen Risiken der KI liefern. Hier bietet sich für die Schweiz die Chance, speziell mit genau diesen Ländern zu KI-Themen zusammenzuarbeiten, insbesondere auch in den Bereichen Entwicklungszusammenarbeit und wissenschaftliche Kooperation. So könnte die Schweiz z.B. Länder wie Kenia dabei unterstützen, den Beobachterstatus bei der KI-Konvention des Europarates zu erlangen und dadurch aktiver an der internationalen KI-Debatte teilzunehmen und globale KI-Governanz mitzugestalten. Oder sie kann durch die DiploFoundation und das Swissnex-Netzwerk den Kompetenzaufbau

zu KI-Themen für Entscheidungsträger:innen im Globalen Süden stärken sowie den Lernerfahrungen aus dem Globalen Süden mehr globale Sichtbarkeit verleihen. Dies auch vor dem Hintergrund des Technologiewettbewerbs von China und den USA, welche diese Länder jeweils umwerben.

4.6 Interner Kulturwandel und Fokussierung auf Trust & Responsible AI angehen

Die Schweiz sollte schlussendlich im Bereich Digitalisierung auf KI-Gouvernanz und Vertrauen im digitalen Raum als Schwerpunktthemen setzen und eine Kompetenzvermutung dahingehend aufbauen. Estland ist ein erfolgreiches Beispiel, wie sich ein kleines Land in einem Themenbereich (*e-Government* bzw. *e-Estonia*) international zu positionieren vermochte. Dabei ist wichtig, eine Positionierung nicht nur gegen aussen anzustreben (siehe Vorschläge 4.1-4.5), sondern die Kompetenz auch im Innern aufzubauen und zu verankern. Vorausschauende Gouvernanz- und Gesetzgebungsprozesse hängen von einem besseren Verständnis der Technologie und ihrer Geschäftsmodelle ab. Wissen zu Themen wie KI gehört zu einem gelebten Selbstverständnis innerhalb der Bundesverwaltung und betrifft alle (*Mindset* und Kultur). Gezielte Schulungen (etwa in Zusammenarbeit mit den technischen Hochschulen ETH und EPFL), Austausch-Programme nach dem Vorbild der *ETH-Fellows* sowie der vermehrte Einbezug einer technologischen Perspektive in unterschiedliche Policy-Felder sind dabei mögliche Massnahmen, wie diese Verankerung gelingen kann. Dafür braucht es jedoch den politischen Willen und die Vision, diesen Themenbereich zu einer politischen Priorität zu machen und höher auf die Agenda zu setzen.

4.7 Schlusswort

Die Debatte um globale KI-Gouvernanz ist derzeit so übersichtlich wie das weltberühmte *Shibuya Scramble Crossing* in Tokio. Gleichzeitig scheint der einzig existierende globale Konsens die Erkenntnis zu sein, dass es Massnahmen braucht, um verantwortungsvolle KI sicherzustellen. Die offene Frage bleibt, welche genau und wie diese umgesetzt werden sollen. Klar ist, es braucht die vertiefte und nachhaltige Auseinandersetzung mit einer Technologie, welche unsere Zukunft noch viel mehr prägen wird als unsere Gegenwart. Das Opportunitätsfenster für die Schweiz, diese Zukunft mitzugestalten, bietet sich jetzt.

Endnoten

- 1 Algorithms and Terrorism: The Malicious use of AI for Terrorists Purposes (2021). United Nations Office of Counter-Terrorism: <https://www.un.org/counterterrorism/sites/www.un.org/counterterrorism/files/malicious-use-of-ai-uncct-unicri-report-hd.pdf>
- 2 Danny Hernandez und Tom B. Brown (2020): Measuring the Algorithmic Efficiency of Neural Network. OpenAI: <https://arxiv.org/ftp/arxiv/papers/2005/2005.04305.pdf>
- 3 Will Henshall (2023): 4 Charts that show why AI progress is unlikely to slow down. Times: <https://time.com/6300942/ai-progress-charts/>
- 4 Sam Jacobs (2023): We chose the Time 100 most influential people in AI. Times: <https://time.com/6311323/how-we-chose-time100-ai/>
- 5 Ronan Farrow (2023): Elon Musk's Shadow Rule. The New Yorker: <https://www.newyorker.com/magazine/2023/08/28/elon-musks-shadow-rule>
- 6 Ian Bremmer and Mustafa Suleyman (2023): The AI Power Paradox. Foreign Affairs: <https://www.foreignaffairs.com/world/artificial-intelligence-power-paradox>
- 7 Cade Metz and Gregory Schmidt (2023): Elon Musk and Others Call for Paus on A.I., Citing "Profound Risks to Society". The New York Times: <https://www.nytimes.com/2023/03/29/technology/ai-artificial-intelligence-musk-risks.html>
- 8 National AI policies & strategies. OECD AI Policy Observatory (besucht am 27.11.2023): <https://oecd.ai/en/dashboards/overview>
- 9 Liebig, L., Güttel, L., Jobin, A. et al. (2022): Subnational AI policy: shaping AI in a multi-level governance system. AI & Soc: <https://doi.org/10.1007/s00146-022-01561-5>
- 10 Gabbott, M. (2023): The Bumpy Road Toward Global AI Governance. Noema Magazine: <https://www.noemamag.com/the-bumpy-road-toward-global-ai-governance/>
- 11 Stefaan G. Verhulst (2023): AI Globalism and AI Localism: Governing AI at the Local Level for Global Benefit. A Response to the On-Going Calls for the Establishment of a Global AI Agency: <https://medium.com/data-policy/ai-globalism-and-ai-localism-governing-ai-at-the-local-level-for-global-benefit-a-response-to-the-b3ac158b72a9>
- 12 G7 Leaders' Statement on the Hiroshima AI Process (2023): <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/g7-leaders-statement-on-the-hiroshma-ai-process/>
- 13 High-level Advisory Body on Artificial Intelligence (2023): United Nations: <https://www.un.org/techenvoy/ai-advisory-body>
- 14 UNESCO (2021): Recommendation on the Ethics of Artificial Intelligence: <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>
- 15 B-Tech Project (2023): OHCHR and business and human rights: <https://www.ohchr.org/en/business-and-human-rights/b-tech-project>
- 16 EU AI Act (2023): first regulation on artificial intelligence: <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>
- 17 Tagesspiegel Background (2023): USA untermauern per Dekret globalen KI-Führungsanspruch. <https://background.tagesspiegel.de/digitalisierung/usa-untermauern-per-dekret-globalen-ki-fuehrungsanspruch>
- 18 The White House (2023): Biden-Harris Administration Secures Voluntary Commitments from Eight Additional Artificial Intelligence Companies to Manage the Risks Posed by AI: <https://www.whitehouse.gov/briefing-room/statements-releases/2023/09/12/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-eight-additional-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/>

- 19 Dan Milmo und Kiran Stacey (2023): Five takeaways from UK's AI safety summit at Bletchley Park. The Guardian: <https://www.theguardian.com/technology/2023/nov/02/five-takeaways-uk-ai-safety-summit-bletchley-park-rishi-sunk>
- 20 Council on Foreign Relations (2023): Cyber Week in Review: <https://www.cfr.org/blog/cyber-week-review-october-20-2023>
- 21 Geneva Internet Platform (2023): ASEAN's innovative AI guide shifts focus, countering EU's regulatory efforts: <https://dig.watch/updates/aseans-innovative-ai-guide-shifts-focus-countering-eus-regulatory-efforts>
- 22 Barbara Grosz (2019): The AI Revolution Needs Expertise in People, Publics, and Societies. MIT Press: <https://hdr.mitpress.mit.edu/pub/wia01ru6/release/5>
- 23 Janis Wong et al. (2022): Key challenges for the participatory governance of AI in public administration. Alan Turing Institute: https://www.turing.ac.uk/sites/default/files/2022-12/key_challenges_for_the_participatory_governance_of_ai_in_public_administration.pdf
- 24 Sam Altman, Greg Brockman and Ilya Sutskever (2023): Governance of superintelligence: <https://openai.com/blog/governance-of-superintelligence>
- 25 Lewis Ho et al. (2023) : International Institutions for Advanced AI: <https://arxiv.org/pdf/2307.04699.pdf>
- 26 State of the Union Address by President von der Leyen (2023): European Commission: https://ec.europa.eu/commission/presscorner/detail/en/speech_23_4426
- 27 Will Henshall (2023): How the U.N. Plans to Shape the Future of AI: <https://time.com/6316503/un-ai-governance-plan-gill/>
- 28 Geoff Mulgan et al. (2023): The world needs a global AI observatory – here's why. MIT Management Sloan School: <https://mitsloan.mit.edu/ideas-made-to-matter/world-needs-a-global-ai-observatory-heres-why>
- 29 Robert F. Trager et al. (2023): International Governance of Civilian AI: A Jurisdictional Certification Approach. Oxford Martin AI Governance Initiative: <https://www.oxfordmartin.ox.ac.uk/publications/international-governance-of-civilian-ai-a-jurisdictional-certification-approach>
- 30 SBFi: Künstliche Intelligenz: Wie geht der Bund damit um? <https://www.sbf.admin.ch/sbf/de/home/dienstleistungen/publikationen/publikationsdatenbank/s-n-2020-6/s-n-2020-6f.html>
- 31 EDA: Strategie Digitalausßenpolitik 2021–2024: https://www.eda.admin.ch/eda/de/home/das_eda/publikationen.html/content/publikationen/de/eda/schweizer-aussenpolitik/Digitalausßenpolitik_2021-2024.html
- 32 Bundesverwaltung: Digitale Schweiz: <https://www.bk.admin.ch/bk/de/home/digitale-transformation-ikt-lenkung/digitale-schweiz.html>
- 33 BAKOM (2022): Monitoring der Leitlinien «Künstliche Intelligenz» für den Bund Evaluation der Anwendung und Aktualität der Leitlinien: <https://cnai.swiss/wp-content/uploads/2022/12/Monitoring-der-Leitlinien-KI-DE.pdf>
- 34 Stellungnahme auf Postulat Dobler 23.320: <https://www.parlament.ch/de/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20233201> und Motion Bendahan
- 35 Technologieneutral bedeutet, dass für ein festgelegtes Ziel, weder der Einsatz einer bestimmten Art von Technologie zu dessen Erreichung vorgeschrieben noch eine Diskriminierung zugunsten dieser Technologie vorgenommen wird.
- 36 EDÖB (2023): Geltendes Datenschutzgesetz ist auf KI direkt anwendbar: https://www.edoeb.admin.ch/edoeb/de/home/kurzmeldung/20231109_ki_dsg.html
- 37 Bundesrat (2023): Bundesrat prüft Regulierungsansätze für Künstliche Intelligenz. Medienmitteilung vom 22. November 2023: <https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-98791.html>

- 38 Swiss Statement (2023): Briefing on Artificial Intelligence: Opportunities and Risks for International Peace and Security: <https://www.aplusforpeace.ch/briefing-artificial-intelligence-opportunities-and-risks-international-peace-and-security#:~:text=In%20addition%2C%20Switzerland%20recently%20launched,to%20peace%20posed%20by%20AI>
- 39 Formal Launch of the Swiss Digital Initiative in Geneva (2020): <https://www.swiss-digital-initiative.org/news/formal-launch-of-the-swiss-digital-initiative-foundation>
- 40 Adrienne Fichter (2023): Wie sich das Internationale Rote Kreuz neu erfindet – und wie die Schweiz dabei verliert. Republik: https://www.republik.ch/2023/07/31/wie-sich-das-internationale-rote-kreuz-neu-erfindet-und-wie-die-schweiz-dabei-verliert?utm_source=newsletter&utm_medium=email&utm_campaign=republik%2Ftem-plate-newsletter-taeglich-2023-07-31
- 41 Innovation Factory AI for Good. ITU (besucht 01.10.23): <https://aiforgood.itu.int/about-ai-for-good/innovation-factory/>
- 42 Kohler, K., Pangrazzi, S. & Zahn, N. (2022): Realitycheck der Ambitionen der Schweizer Digitalausserpolitik. Zürich: foraus - Forum Aussenpolitik.
- 43 EDA: Strategie Digitalausserpolitik 2021–2024. https://www.eda.admin.ch/eda/de/home/das_eda/publikationen.html/content/publikationen/de/eda/schweizer-aussenpolitik/Digitalausserpolitik_2021-2024.html
- 44 International Atomic Energy Agency
- 45 International Plant Protection Convention
- 46 MIT (2023): The world needs a global AI observatory — here's why: <https://mitsloan.mit.edu/ideas-made-to-matter/world-needs-a-global-ai-observatory-heres-why>
- 47 BBI 2019 2363: Bundesbeschluss über die Massnahmen zur Stärkung der Rolle der Schweiz als Gaststaat 2020–2023: <https://www.fedlex.admin.ch/eli/fga/2019/669/de>
- 48 Brian O’Gara (2020): Citizens’ Assemblies: How Deliberative Democracy Can Help Control Artificial Intelligence. In: The Human Right to Democratic Control of Artificial Intelligence: https://humanrights.gov.au/sites/default/files/2020-07/70_-_jonathan_crock_et_al_1.pdf
- 49 Michele Gilman (2023): Democratizing AI: Principles for Meaningful Public Participation. University of Baltimore School of Law: https://datasociety.net/wp-content/uploads/2023/09/DS_Democratizing-AI-Public-Participation-Brief_9.2023.pdf

Zitierweise

Eggler, S. und Paeffgen, N. (2023). Globale Gouvernanz Künstlicher Intelligenz und die Rolle der Schweiz. Bern: foraus - Forum Aussenpolitik.

Danksagung

Im Schreibprozess dieses *Policy Papers* durften die Autorinnen auf die Erfahrung und umfassende Expertise einer Vielzahl von engagierten Personen und Expert:innen zurückgreifen. Ein spezieller Dank geht an Dr. Isabel Ebert, Dr. Anna Jobin, Kathleen Ziemann, Nicolas Zahn, Oliver Guggenbühl sowie das foraus Review-Board, insbesondere an Oskar Jönsson, Selina Morell und David Svarin, und dem foraus Vorstand. Sie unterstützten die Autorenschaft durch wertvolle Gedanken, kritische Rückmeldungen und motivierende Worte im Schreib- und Review-Prozess. Besonderer Dank geht ebenfalls an das gesamte Team der foraus Geschäftsstelle rund um die Co-Leitung durch Seraina Capatt und Darius Farman sowie das Kommunikationsteam um Yevgeniya Frei.

Disclaimer

Die vorliegende Publikation gibt die persönliche Meinung der Autor:innen wieder und entspricht nicht zwingend derjenigen des Vereins foraus, der externen Reviewer:innen oder der Arbeitgeber der Autorinnen.

www.foraus.ch

Bern | foraus – Forum Aussenpolitik | Marktgasse 36 | 3011 Bern
office@foraus.ch | +41 44 501 68 65

Genève | foraus – Forum de politique étrangère | c/o Impact Hub, Rue de Fendt 1 |
1201 Genève | bureau_romandie@foraus.ch | +41 22 559 96 96

IBAN: CH06 0900 0000 6017 6892 9

Bei foraus aktiv werden

Durch das einzigartige Grassroots-Modell des Vereins setzt sich foraus für eine gesellschaftspolitische Zukunft der Schweiz ein und steht für eine konstruktive Aussenpolitik und einen informierten Dialog. Du teilst diese Vision? Werde jetzt bei uns aktiv und unterstütze uns! ●

als Mitglied

Eine Mitgliedschaft in unserem einzigartigen Netzwerk und ein ehrenamtliches Engagement bei foraus stehen jeder und jedem offen. Wir bieten dir Zugang zu einem hochkarätigen Netzwerk, spannenden Persönlichkeiten der Schweizer Aussenpolitik und der Möglichkeit, dein wissenschaftliches Know-How in die öffentliche Debatte zu tragen.

als Autor:in

foraus ermöglicht es dir, Herausforderungen der Schweizer Aussenpolitik konkret anzupacken und bietet dir eine Plattform dafür, deine innovativen Ideen für die Schweizer Aussenpolitik im Rahmen eines Diskussionspapiers oder einer Kurzanalyse zu publizieren.

als Gönner:in

Unser Gönnerverein «Cercle des donatrices & donateurs» trägt zur Verbreiterung der Trägerschaft bei und bietet interessierten Persönlichkeiten die Möglichkeit, foraus nachhaltig zu unterstützen und zu fördern.



foraus.ch/mitglied-werden