Policy Perspectives
Vol. 7/2, März 2019

CSS
ETH Zurich

Ein neutraler Hub für KI-Forschung

Im internationalen Politikdiskurs wird Künstliche Intelligenz (KI) oft auf einen Technologiewettlauf verengt. Internationale Wissenschaftsdiplomatie kann indes die Chancen der neuen Technologien fördern und die Risiken frühzeitig mindern.

Von Sophie-Charlotte Fischer und Andreas Wenger

Künstliche Intelligenz wirkt als Treiber eines umfassenden Wandels von Wirtschaft, Gesellschaft und Staat. Obwohl der Beginn der KI-Forschung bis in die 1950er Jahre zurückreicht, ist KI erst in den letzten Jahren zu einem prägenden Thema auf der internationalen politischen Agenda geworden. Die Ursachen dafür sind primär technischer und wirtschaftlicher Natur. Die Kombination von drei Faktoren – schnell steigende Rechenleistungen von Computern, wachsende Datenmengen und optimierte maschinelle Lernansätze und Algorithmen – hat während der letzten 15 Jahre eine neue Welle des Fortschritts in KI ermöglicht. KI-Technologien aus Laboren von Unternehmen und Universitäten finden Eingang in immer mehr praktische Verwendungen. Dabei zeigt sich, dass mit KI sowohl grosse wirtschaftliche und gesellschaftliche Chancen als auch beträchtliche Risiken verbunden sind.

Grosse globale Technologieunternehmen, insbesondere in den USA und in China, verfügen heute über einen Wissens-, Daten- und Anwendungsvorsprung im Bereich der KI. Gesellschaftliche und staatliche Akteure weisen dagegen einen Aufholbedarf auf. Die Führungsrolle der Technologieunternehmen ergibt sich aus den exzellenten Rahmenbedingungen, welche sie für die Entwicklung und Kommerzialisierung von KI bieten. Diese umfassen grosse Rechenleistung, vielfältige Datensätze und beträchtliche finanzielle Investitionen in Forschung und Entwicklung. In Kombination mit attraktiven Löhnen, flexiblen Arbeitsbedingungen und Freiraum für kreatives Arbeiten an interessanten Problemstellungen erlaubt dies den Firmen, die besten KI-Forscher aus aller Welt anzuwerben. Gleichzeitig wirkt sich die Konzentration von KI-Forschung und Entwicklung in wenigen grossen Unternehmen zunehmend negativ auf die Konkurrenzfähigkeit öffentlicher Forschungsinstitutionen sowie kleinerer Firmen aus.

Staaten wiederum antizipieren umfassende Auswirkungen von KI – als «Dual Use»-Technologie – auf das

Schlüsselaussagen

- Staatliche Akteure sehen KI zunehmend als strategische Ressourcen an. Sie nehmen vermehrt Einfluss auf den Innovationsprozess und die Weiterverbreitung von KI-Technologien.
- Im Kontext einer wachsenden Wettbewerbsrhetorik steigen die mit unausgereiften KI-Technologien verbundenen Risiken.
- Erforderlich ist ein politisch neutraler Hub für KI-Forschung, der einer verantwortungsvollen, integrativen und friedlichen Entwicklung und Nutzung der neuen Technologien verpflichtet ist.
- Die Schweiz verfügt über ein dynamisches Ökosystem im KI-Bereich und ist auch politisch gut positioniert, um als Gaststaat für KI-Governanz-Initiativen zu wirken.

weltweite wirtschaftliche und militärische Machtgefüge. KI wird in wachsendem Masse als strategische Ressource gesehen und behandelt. Immer mehr Staaten erarbeiten nationale KI-Strategien mit dem Ziel, sich eine vorteilhafte Ausgangsposition im KI-Bereich zu verschaffen. Aufgrund des grossen wirtschaftlichen Potentials und der sicherheitspolitischen Relevanz von KI versuchen Staaten zunehmend, Einfluss auf den Innovationsprozess und die Weiterverbreitung der neuen Technologien zu nehmen. Die Ausbildung von KI-Talenten sowie die Forschung und Entwicklung werden gezielt gefördert. In Partnerschaften mit Technologiekonzernen soll KI für öffentliche Zwecke akquiriert und genutzt werden. Gleichzeitig versuchen Regierungen, ihre nationalen KI-Ressourcen gegenüber unerwünschten internationalen Einwirkungen abzuschotten und den Transfer von KI-Technologien an Wettbewerber zu verhindern, auch wenn der Effekt solcher Maßnahmen umstritten ist. Beispielhaft hierfür sind Bemühungen zum Schutz der nationalen Technologie- und Industriebasis vor ausländischen Direktinvestitionen sowie Erwägungen, Exporte von spezifischen KI-Technologien an Wettbewerber einzuschränken.

Die Konzentration von KI-Ressourcen in privaten Technologieunternehmen und die damit verbundene politische Wettbewerbsdynamik zwischen den Grossmächten stellen KI-schwache Gesellschaften und Regionen vor schwierige Herausforderungen. Wirtschaftlich drohen sie zunehmend verwundbar gegenüber globalen technologischen Oligopolen zu werden; politisch könnten sie mehr und mehr abhängig von den politischen Entscheidungen anderer Akteure werden. Gleichzeitig besteht die Gefahr, dass die mit der Entwicklung von KI verbundenen grundsätzlichen technologischen und gesellschaftlichen Risiken in den Hintergrund treten. Im Kontext eines technologischen Wettlaufs kann die Wahrnehmung, dass in strategisch wichtigen Bereichen ein «first-mover advantage» existiert, eine Abwärtsspirale auslösen. Wenn immer weniger Akteure unter immer grösseren Zeitdruck über die Zukunft einer transformativen Technologie entscheiden, dann steigt auch die Wahrscheinlichkeit von technischen Unfällen aufgrund unausgereifter KI-Technologien.

Diese Dynamiken bergen das Risiko, dass die bereits bestehenden Ungleichheiten zwischen Gesellschaften und Regionen weiter verstärkt werden. Der Prozess der Entwicklung von neuen Technologien sollte aber nicht von wenigen dominanten wirtschaftlichen und politischen Akteuren geprägt sein. In einer globalisierten Welt braucht es die Einbindung von wirtschaftlichen, gesellschaftlichen und politischen Akteuren aus allen Regionen mit dem Ziel, die mit KI verbundenen Chancen allen zugänglich zu machen. Es ist notwendig, vorausschauend effektive Kanäle für internationale Kooperation zu schaffen, um die mittel- und langfristige Entwicklung und Nutzung von KI transparent, fair und nachhaltig zu gestalten.

Wissenschaftsdiplomatie ermöglicht KI-Zusammenarbeit

Die Verbindung von Wissenschaft und Diplomatie stellt ein vielversprechendes Instrument dar, um den mit KI verbundenen technischen und gesellschaftlichen Risiken frühzeitig entgegenzuwirken. Die Wissenschaftsdiplomatie eignet sich, um Brücken zwischen Gesellschaften zu bauen und gemeinsame Strategien zur Überwindung internationaler Herausforderungen zu entwickeln. Der wissenschaftliche Austausch ermöglicht einen konstruktiven Dialog zwischen gesellschaftlichen, wirtschaftlichen und staatlichen Akteuren, weil er sich an den Werten und methodischen Standards der Wissenschaft orientiert. Forschern geht es in ihrer Arbeit um Kooperationen zugunsten neuer wissenschaftlicher Erkenntnisse. Der wissenschaftliche Austausch unterstützt die Verständigung und Etablierung nachhaltiger persönlicher und institutioneller Verbindungen über politische und ideologische Grenzen hinweg.

Der Blick in die Vergangenheit zeigt, dass Wissenschaftsdiplomatie bereits früher einen entscheidenden Beitrag zum gegenseitigen Verständnis in den internationalen Beziehungen geleistet hat. Wissenschaftliche Megaprojekte wie die Internationale Raumstation oder das CERN (The European Organization for Nuclear Research) haben sich als effektive Mechanismen erwiesen,

Weiterführende Literatur

AI Governance: A Research Agenda *Dafoe, A., Center for the Governance of Artificial Intelligence, University of Oxford, 2017.*
Dieser Report skizziert ein umfassendes neues Forschungsprogramm in KI-Gouvernanz.

Concrete Problems in AI Safety. *Dario Amodei, Chris Olah, Jacob Steinhardt, Paul Christiano, John Schulman, Dan Mané, 2016.*
Dieser Artikel untersucht verschiedene Herausforderung in der Sicherstellung, dass moderne Machine Learning Systeme wie intendiert funktionieren und Unfälle vermieden werden können.

The CERN Community; A Mechanism for Effective Global Collaboration? *Mark Robinson, Global Policy, vol.10 (1), 41–51, 2018.*
Dieser Artikel untersucht am Beispiel des CERN, warum globale wissenschaftliche Megaprojekte besonders effektiv in der Ermöglichung internationaler Kollaboration sind.

um Verhärtung und Stillstand in der internationalen politischen Zusammenarbeit entgegenzuwirken.

Die Wissenschaftsdiplomatie als ein Lösungsansatz für globale KI-Herausforderungen erscheint auch daher aussichtsreich, weil die KI-Forschungsgemeinschaft sehr international geprägt ist und eine aussergewöhnlich offene Kultur etabliert hat. Viele KI-Forscher veröffentlichen ihre Resultate auf der Onlineplattform arXiv, um sie für andere interessierte Nutzer frei zugänglich zu machen und sie so weiterzuentwickeln. Staatliche Bestrebungen, länderübergreifende Forschung und den wissenschaftlichen Austausch im KI-Bereich einzuschränken, haben heftige Kritik von Seiten der Wissenschaft ausgelöst.

Ein neutraler Hub für KI-Forschung

Als ein Element einer globalen KI-Gouvernanz-Architektur schlagen wir die Schaffung eines politisch neutralen, internationalen und interdisziplinären Hubs für KI-Grundlagenforschung vor, der einer verantwortungsvollen, integrativen und friedlichen Entwicklung und Nutzung von KI verpflichtet ist. Die Verknüpfung zwischen Wissenschaft und Diplomatie soll dazu beitragen, dass ein nachhaltiger Mechanismus für internationale Kooperation im Bereich KI geschaffen wird und das Potential von KI einem möglichst grossen Teil der Weltbevölkerung zugänglich gemacht werden kann. Die Lancierung einer solchen Initiative würde signalisieren, dass die Entwicklung von KI nicht zwangsläufig verläuft, sondern in einem internationalen Rahmen gestaltet werden kann. Ferner würde die Initiative die Überzeugung stützen, wonach es sich bei der Entwicklung und Anwendung von KI nicht um ein Nullsummenspiel handelt. Die Vision einer globalen Forschungsplattform im Bereich KI ist bereits in verschiedenen Kontexten zur Sprache gekommen, so beispielsweise am AI for Good Summit oder auch in Diskussionen im Rahmen der OECD. Im Folgenden geht es darum, eine mögliche Version einer solchen Initiative zu konkretisieren und eine öffentliche Auseinandersetzung über die Chancen und Herausforderungen sowie eine mögliche Ausgestaltung voranzutreiben. Weiterhin wird diskutiert, welche besondere Rolle die Schweiz in der Umsetzung eines solchen Vorhabens spielen könnte.

Ein internationaler Hub für KI Grundlagenforschung hätte vier zentrale Funktionen, um einigen der sichtbar werdenden Spannungsfelder auf internationaler Ebene entgegenzuwirken:

Erstens soll eine weltweit kompetitive, internationale und integrative Plattform für KI-Grundlagenforschung aufgebaut werden. Der Hub soll als attraktives Laboratorium für die besten KI-Talente aus aller Welt dienen, um die Grundlagen für transparente KI-Anwendungen über Landesgrenzen hinweg zu schaffen und KI als globales Kollektivgut zu entwickeln. Der Forschungsrahmen der Plattform soll ein breites Spektrum von methodischen Ansätzen in der KI-Forschung abdecken. Auch wenn ein Grossteil der aktuellen KI-Fortschritte auf sogenannten «Deep Learning»-Algorithmen beruht, sollen auch alternative Paradigmen sowie Forscher verwandter Disziplinen, wie beispielsweise den Neurowissenschaften, einbezogen werden, da sie interessante neue Impulse in der KI-Grundlagenforschung geben können. Die Initiative soll sich an den höchsten wissenschaftlichen und ethischen Standards orientieren und ein attraktives Arbeitsumfeld für die besten KI-Talente bieten.

Zweitens soll der KI-Hub als Plattform dienen, um mit KI verbundene technische und gesellschaftliche Risiken zu erforschen, zu reflektieren und zu bewältigen. Dazu zählen beispielsweise mit KI verbundene technische Sicherheitsrisiken sowie die Nachvollziehbarkeit algorithmischer Entscheidungsfindungsprozesse. Zur Erforschung dieser Risiken sollen ausserdem Wissenschaftler aus den Human- und Sozialwissenschaften einbezogen werden.

Drittens soll der KI-Hub einen Beitrag zur Entwicklung von Normen und «best practices» hinsichtlich der Nutzung von KI leisten. Die internationale interdisziplinäre Forschungsgemeinschaft soll sich beratend in künftige KI-Gouvernanzprozesse und -Foren einbringen. Somit könnte einer Prägung von KI-Normen durch wenige mächtige Akteure entgegengewirkt werden.

Viertens soll die Initiative als Aus- und Weiterbildungszentrum für KI-Forscher dienen. Master- und PhD-Studenten, die einen Teil ihrer Ausbildung am KI-Hub absolvieren, sollen die Rolle von Multiplikatoren übernehmen, wenn sie ihre Karriere an anderen Organisationen fortsetzen. Über Forschungs- und Ausbildungspartnerschaften trägt der KI-Hub zur Stärkung der erkenntnisorientierten Gemeinschaft in diesem Bereich bei. Eine zusätzliche Möglichkeit ist, dass sich Forscher verschiedener Disziplinen und Nationalitäten an dem Zentrum flexibel in Projektgruppen zusammenschliessen, um gezielt mit KI-Anwendungen zur Lösung globaler Probleme beizutragen, wie sie beispielsweise in den United Nations Sustainable Development Goals formuliert werden. Denkbar wäre hier auch ein dezentraler Ansatz, der die Etablierung projektbasierter Ableger des Hubs in anderen Weltregionen einschliesst, um so problemorientierte KI-Anwendungen mit lokaler Expertise und betroffenen Gesellschaftsgruppen dynamisch entwickeln zu können.

Gouvernanz des KI-Hubs

Die Mitgliedschaft in diesem internationalen KI-Hub soll grundsätzlich allen Staaten offenstehen. Es erscheint jedoch zweckmässig, dass das wissenschaftliche Grossprojekt von einer «Koalition der Willigen» angestossen wird. Wichtig ist von Beginn an das Engagement einer überregionalen Gruppe gleichgesinnter Staaten. Damit soll vermieden werden, dass die Initiative als Vorhaben einer selektiven Staatengruppe wahrgenommen wird, die Wettbewerbsfähigkeit ihrer Region einseitig voranzutreiben, und andere Länder abschreckt, ebenfalls Teil des Projekts zu werden. Der KI-Hub erfüllt dann seinen Zweck, wenn er von der internationalen Expertengemeinschaft

akzeptiert wird und einen wichtigen Knotenpunkt in einem globalen Netzwerk für sichere und verantwortungsvolle KI darstellt.

Die am Forschungshub gewonnen Ergebnisse sollen den Mitgliedsstaaten als Kollektivgut zur Verfügung gestellt werden. Ähnlich wie beim CERN sollen die Forschungsgruppen des KI-Hubs weltweit mit Partnerlaboren kooperieren können; dies allerdings ebenfalls nur dann, wenn die daraus hervorgehenden Forschungsresultate allen anderen Mitgliedsstaaten zugänglich gemacht werden. Alle Mitgliedsstaaten sollen zudem von einer gemeinsam geschaffenen, exzellenten Infrastruktur für KI-Forschung profitieren, die einzelne Länder im Alleingang nicht erschaffen können. Finanzielle Beiträge der Mitglieder sollen sich an den jeweiligen wirtschaftlichen Möglichkeiten orientieren und zusätzlich die Form von «in kind contributions» annehmen können. Damit soll die Schwelle für eine Mitgliedschaft für wirtschaftlich schwächere Staaten gesenkt werden.

Es sind verschiedene Modelle zur Anbindung des KI-Hubs an die Uno denkbar. Der Hub könnte ähnlich der IAEO (Internationale Atomenergie-Organisation) als autonome wissenschaftlich-technische Organisation etabliert werden, die über ein Kooperationsabkommen mit der Uno verbunden ist und regelmässig an die Generalversammlung und den Sicherheitsrat der Uno berichtet.

Eine Rolle für die Schweiz

Die Schweiz ist aus verschiedenen Gründen besonders gut geeignet, um die Vision eines internationalen KI-Forschungshubs in einer führenden Rolle voranzutreiben und in der Umsetzung allenfalls als Gaststaat zu wirken. Als Uno-Standort, Nicht-Mitglied der EU, und als eines der am stärksten globalisierten Länder weltweit könnte die Schweiz eine wichtige Brückenfunktion übernehmen, um eine überregionale Ausrichtung des Hubs von Beginn an sicherzustellen. Ihre politische Neutralität, Stabilität und Eigenständigkeit sowie ihre Erfahrung bei der Mitgestaltung von Multi-Stakeholder-Prozessen machen die Schweiz zu einem glaubwürdigen Gastgeber einer globalen Forschungsplattform für KI. Zu dieser Wahrnehmung trägt ebenfalls der besondere historische Einsatz der Schweiz dafür bei, dass am CERN ausschliesslich zivile und keine militärische Forschung betrieben wird.

Die Schweiz bietet zudem eine vorteilhafte Kombination wirtschaftlicher und wissenschaftlicher Voraussetzungen, die für die Umsetzung eines solchen Projekts essentiell sind. Das Land verfügt bereits über ein sehr dynamisches Ökosystem im KI-Bereich und in angrenzenden Disziplinen, das exzellente technische Universitäten wie die ETH und EPFL, eine lebendige Startup-Szene sowie global führende Technologieunternehmen einschliesst. Mit der gemeinsamen Beteiligung von Bund, Forschung und Industrie an der Initiative würden die nationalen Kompetenzen besser gebündelt und aufgrund ihrer kritischen Masse auch international noch besser vernetzt. Als Standort des KI-Hubs würde die Schweiz durch die hohe Lebensqualität und die zentrale Lage im Herzen Europas eine attraktive Alternative zu Orten wie dem Silicon Valley und anderen bevorzugten Destinationen der weltweit besten KI-Forscher bieten.

Auf internationaler Ebene bietet sich der Schweiz die Chance, sich als führender globaler Forschungs- und Innovationsstandort, als wichtiger Impulsgeber der internationalen Zusammenarbeit sowie als Brückenbauer an den Schnittstellen zwischen Friedenspolitik und Technologieaussenpolitik in Erinnerung zu rufen. Innenpolitisch bietet die Initiative für die stark vernetzte und vom Aussenhandel abhängige Schweiz die Chance, einen glaubwürdigen Beitrag zur Entwicklung globaler KI-Normen zu leisten. Der Beitrag soll die Werte und Interessen der Schweiz wiederspiegeln und das Wissen und die Kompetenzen einer innovativen Verwaltung, einer agilen Industrie und einer exzellenten Forschung bündeln. Der Einsatz zugunsten eines politisch neutralen, internationalen und interdisziplinären Hubs für KI ist richtungsweisend, weil es das in der eigenen Geschichte verankerte Engagement der Schweiz für den Frieden mit der Technologiekompetenz eines hochentwickelten Kleinstaates kombiniert.

Sophie-Charlotte Fischer ist Doktorandin am Center for Security Studies (CSS) an der ETH Zürich.

Andreas Wenger ist Professor für Internationale und Schweizer Sicherheitspolitik an der ETH Zürich und Direktor des Center for Security Studies (CSS).

Policy Perspectives is edited by the Center for Security Studies (CSS) at ETH Zurich. The CSS is a center of competence for Swiss and international security policy. It offers security policy expertise in research, teaching, and consultancy. The CSS promotes understanding of security policy challenges as a contribution to a more peaceful world. Its work is independent, practice-relevant, and based on a sound academic footing.

Editor: John (Jack) Thompson

Comments welcome at PolicyPerspectives@sipo.gess.ethz.ch

Recent editions of **Policy Perspectives**:
- **Trump's Missile Defense: Challenges for Europe** (7/1) by Oliver Thränert
- **The Use of Smart Pressure to Resolve Civil Wars** (6/8) by Dr. Allard Duursma

For more editions, and for a free online subscription, visit www.css.ethz.ch/en/publications/css-policy-perspectives

© 2019 Center for Security Studies (CSS), ETH Zurich www.css.ethz.ch

ISSN: 2296-6471; DOI: 10.3929/ethz-b-000332541