Warum Deutschland ein AI Security Institute braucht

Philip Fox und Anton Leicht, Kira Center

erschienen im Tagesspiegel Background Digitalisierung & KI, 17.03.2025

Immer mehr Länder haben KI-Sicherheitsinstitute gegründet. Nur Deutschland und Italien bleiben im internationalen Netzwerk außen vor. Angesichts der geopolitischen Entwicklungen ist es umso wichtiger, endlich nachzuziehen, schreiben Philip Fox und Anton Leicht vom Thinktank Kira.

Zwischen den führenden KI-Nationen der Welt hat spätestens im vergangenen Jahr ein Wettlauf begonnen, nur Deutschland ist noch nicht einmal im Stadion. Sowohl die USA als auch China wetten weiterhin auf große Fortschritte in der KI-Forschung: Sie gehen erstens davon aus, dass KI-Systeme in wenigen Jahren eine absolut zentrale wirtschaftliche und gesellschaftliche Bedeutung haben werden. Und zweitens, dass KI eine „Dual Use"-Technologie ist, die nicht nur gesellschaftlichen Nutzen stiftet, sondern auch eine militärische Gefahr darstellt, wenn sie von geopolitischen Rivalen eingesetzt wird. Beide Länder arbeiten deshalb intensiv an staatlichen Strategien, um heimische KI-Firmen zu fördern und gefährliche KI-Entwicklung im In- und Ausland zu überwachen und zu kontrollieren, beispielsweise durch Exportverbote.

Geht die Wette der globalen Großmächte auf, ist Deutschland schnell außen vor. In den chinesischen Plänen sind wir allenfalls peripherer Gegner; in den US-Plänen sind wir, gerade unter Trump, auch nicht unbedingt ein essenzieller Partner. Ein viel beachteter Bericht der Hoover Institution, einem einflussreichen konservativen Thinktank, schlägt zum Beispiel eine US-geführte Koalition der Five-Eyes-Staaten vor (USA, Großbritannien, Kanada, Australien, Neuseeland). Eine Teilhabe Deutschlands an gemeinsamen Defensivprojekten, zum Beispiel im Bereich Cybersicherheit oder der ökonomischen Nutzung von Spitzenmodellen, ist alles andere als eine Selbstverständlichkeit.

Dieses Wettrennen ist aktuell kein Gegenstand der deutschen Politik. KI taucht zwar als Begriff auf – ein bisschen „Schlüsseltechnologie" hier, ein bisschen „KI-Strategie" dort. Aber nichts davon wird dem disruptiven Potenzial von KI und auch ihrer gravierenden Bedeutung für die nationale Sicherheit gerecht.

Großbritannien hat es vorgemacht

Andernorts werden Fakten geschaffen: In einer einzigartigen Aktion hat das Vereinigte Königreich 2023 das weltweit erste „AI Security Institute" (Aisi) ins Leben gerufen, das mittlerweile über 100 Angestellte hat (bei der Gründung hieß es „AI Safety Institute" und wurde im Februar in „Security Institute" umbenannt). Ausgestattet mit genügend Ressourcen rekrutierte das Aisi binnen kürzester Zeit angesehene, internationale Fachleute und entwickelte sich in ein global anerkanntes Kompetenzzentrum. Heute ist Großbritannien auch dank des erfolgreichen Aisi die drittwichtigste Kraft in der globalen KI-Politik, hinter den USA und China.

Deutschland sollte so schnell wie möglich mit einem eigenen Aisi nach internationalem Vorbild nachziehen. Ein KI-Sicherheits-Institut wäre für die Bundesregierung die primäre Anlaufstelle und Beratungsinstanz hinsichtlich einer zukunftsgerichteten und evidenzbasierten Perspektive auf Spitzen-KI. Ihr enges Mandat wäre auf höchstens drei Kernziele beschränkt:

1. Früherkennung neuer KI-Trends und ihrer politischen Implikationen
2. Beratung der Bundesregierung in Fragen von KI und nationaler Sicherheit
3. technische Forschung zu KI-gestützten Bedrohungen für die nationale Sicherheit

Ziel eines Aisi ist es nicht, ganz allgemein die gesellschaftlichen Folgen von KI in den Blick zu nehmen oder parallel zur europäischen KI-Verordnung neue Regulierung zu schaffen. Vielmehr geht es ausschließlich um den Aufbau dringend benötigter staatlicher Kapazität, um Risiken für die nationale Sicherheit zu minimieren und geopolitische Souveränität angesichts der rasanten KI-Entwicklung sicherzustellen.

Immer mehr Länder haben AI Safety Institute

Viele andere Länder haben den Bedarf für AI Safety oder Security Institute erkannt, darunter neben Großbritannien auch Japan, Kanada, Indien und die USA. Frankreich hat im Februar die Gründung eines eigenen Sicherheitsinstituts bekannt gegeben. Aisis liefern schon nach kurzer Zeit greifbare Ergebnisse: Das amerikanische und britische Aisi testen Spitzenmodelle vor deren Veröffentlichung; das britische Institut hat sich durch seinen engen Fokus auf Fragen der nationalen Sicherheit schon nach kurzer Zeit als führend in Teilbereichen der technischen Risikoforschung etabliert. Die einzelnen Institute haben sich zudem in einem internationalen Netzwerk zusammengeschlossen, das angesichts der grenzüberschreitenden Auswirkungen von KI auf Zusammenarbeit in Forschung und Risikosteuerung setzt.

Wer entgegen ursprünglicher Pläne nicht Teil des Netzwerks ist: Deutschland. Eine indirekte Beteiligung gibt es allenfalls über das europäische AI Office, das als Regulierungsbehörde eine nicht ganz einfache Stellung im Netzwerk einnimmt. Deutschland fällt hier nicht nur hinter die eigenen Ansprüchen zurück, sondern auch hinter andere europäische Länder wie Frankreich oder Großbritannien.

Ein deutsches Aisi ist nicht bloß nötig, um den Anschluss an den internationalen KI-politischen Diskurs zu behalten und mit am Tisch zu sitzen, wenn einflussreiche Staaten über eine der zentralen Technologien der Zukunft diskutieren. Vielmehr sind beim Thema KI-Sicherheit vitale Interessen der Bevölkerung betroffen, die auf oberster Regierungsebene geschützt werden müssen. Dazu braucht es heimische, souveräne Expertise, die ein deutsches Aisi aufbauen und bündeln würde. Andernfalls müssten wir uns zu sehr auf die Bereitschaft von Partnerländern wie den USA oder Großbritannien verlassen, sicherheitskritische Erkenntnisse ihrer Aisis rechtzeitig mit Deutschland zu teilen. Das könnte sich als frommer Wunsch erweisen – insbesondere dann, wenn diese Erkenntnisse primär die nationalen Interessen dieser Länder betreffen. Damit Informationen überhaupt sinnvoll geteilt werden können, müssten ausländische Organisationen zudem sicher sein, dass es in der deutschen Regierung einen Ansprechpartner mit der nötigen Expertise gibt, um diese Informationen zu bewerten und daraus sinnvolle Schritte abzuleiten.

Genauso wenig reicht es, sich allein auf das European AI Office zu verlassen. Dieses hat als EU-Behörde kein Mandat für Fragen der nationalen Sicherheit und als Regulierungsbehörde ohnehin eine andere Funktion. Ein beratendes Aisi, das nicht gleichzeitig reguliert, könnte zu KI-Unternehmen leichter die vertrauensvolle Arbeitsbeziehung aufbauen, die für eine Zusammenarbeit in sicherheitskritischen Fragen unerlässlich ist.

Arbeitsteilung zwischen AISI und BNetzA

Aus demselben Grund braucht es eine klare Arbeitsteilung zwischen einem deutschen Aisi und der Bundesnetzagentur (BNetzA). Letztere hat als geplante designierte Behörde zur Umsetzung der europäischen KI-Verordnung die Funktion der Marktaufsicht. Ein Aisi sollte sich hingegen auf ein wissenschaftlich und politisch unabhängiges Beratungsmandat mit engem Fokus auf Risiken für die nationale Sicherheit beschränken, das nicht im Aufgaben- oder Kompetenzbereich der BNetzA liegt.

Auch ein kürzlich erschienener offener Brief aus der deutschen Wissenschaft ruft zur Gründung eines Aisi auf. Darin weisen renommierte KI-Fachleute darauf hin, dass während international schon die „zweite Welle“ an Aisis entsteht, in Deutschland nach wie vor ein „gemeinsames und koordiniertes Vorgehen“ fehlt. Entscheidend für den Erfolg eines deutschen Aisi wird aus unserer Sicht vor allem die konkrete Umsetzung. Ein Aisi muss schnell und unbürokratisch auf dringliche Entwicklungen reagieren können und in der Lage sein, (inter-)nationale Top-Talente zu wettbewerbsfähigen Bedingungen zu rekrutieren.

Wir schlagen deshalb vor, dem erfolgreichen Beispiel Großbritanniens zu folgen und ein Aisi nicht als klassische Behörde oder universitäre Einrichtung zu strukturieren. Im deutschen Kontext bietet sich dafür die Rechtsform der bundeseigenen GmbH an. Das Beispiel der Bundesagentur für Sprunginnovationen (Sprind) zeigt, wie sich auf diese Weise agile, staatliche Institutionen mit Start-up-Charakter aufbauen lassen. Ein enger Fokus auf Risiken für die nationale Sicherheit sichert zudem die Handlungsfähigkeit. Im Bereich Sicherheitsstandards sollte sich ein AISI auf technische Forschung beschränken, um keine Doppelstrukturen zu Normierungs- oder Regulierungsbehörden zu schaffen.

Deutschland und Italien fehlen im internationalen Netzwerk

Ohnehin hatte Deutschland schon 2024 in einer gemeinsamen Erklärung mit neun anderen Ländern und der EU empfohlen, dass Staaten eigene Institute für KI-Sicherheit gründen sollten. Passiert ist seitdem nichts. In Ermangelung eines eigenen Aisi ist Deutschland neben Italien der einzige Unterzeichner, der heute nicht Teil des internationalen Netzwerks der KI-Sicherheitsinstitute ist. Die neue Bundesregierung sollte mit hoher Priorität dafür sorgen, dass Deutschland eigenständiger Teil dieses wichtigen internationalen Kreises wird. Genügend Ressourcen für ein konkurrenzfähiges Aisi hätte Deutschland in jedem Fall. Zum Vergleich: Großbritannien rechnet mit circa 60 Millionen Euro jährlich in den kommenden Jahren. Ein Budget in dieser Größenordnung wäre ein Zeichen, dass Deutschland die sicherheitspolitische Dimension von KI ernst nimmt.

Weltweit wird diese Dimension zunehmend anerkannt. Deutschland sollte jetzt nachziehen und mit einem Aisi nach internationalem Vorbild mehr staatliche Kapazität für das Thema schaffen. Die neue Bundesregierung könnte dann jederzeit auf wissenschaftliche Top-Expertise zurückgreifen, um auf die weitreichenden Auswirkungen von KI auf die nationale Sicherheit vorbereitet zu sein.