Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement für Verteidigung
Bevölkerungsschutz und Sport
**Generalsekretariat VBS**
Sicherheitspolitik

# Künstliche Intelligenz in der Cybersicherheit und Sicherheitspolitik

## Bericht der Projektgruppe «Künstliche Intelligenz in der Cybersicherheit und Sicherheitspolitik»

Dieser Fachbericht wurde im Rahmen der Arbeiten der interdepartementalen Arbeitsgruppe «Künstliche Intelligenz» von der Projektgruppe «Cybersicherheit und Sicherheitspolitik» unter Co-Federführung des EDA (Büro Sondergesandter für Cyber-Aussen- und Sicherheitspolitik) und des VBS (Sicherheitspolitik – Cyberdefence) verfasst.

**Mitglieder der interdepartementalen Projektgruppe "*Cybersicherheit und Sicherheitspolitik*"**

Koordination:
- Maurice Eglin (GS-VBS – Cyberdefence; Leiter Aussen- und Krisenkoordination; Ausbildung)

Projektgruppe:
- Dr. Albert Blarer (VBS; armasuisse W+T; Wissenschaftlicher Projektleiter)
- Dr. Stefan Brem (VBS; BABS; Chef FB Risikogrundlagen und Forschungs-Koordination / Stv. C GB)
- Sandra Caluori (EDA; DV, Sektion Völkerrecht)
- Dr. Laura Crespo (VBS; NDB; Steuerungsverantwortliche Cyber)
- Dr. Jon Fanzun (EDA; Sondergesandter; Chef des Büros für Cyber-Aussen und Sicherheitspolitik)
- Oberst i Gst Robert Flück (VBS; FUB; Projektleiter Cyber der Armee)
- Patrick Gansner (GS VBS – Sicherheitspolitik; Chef Strategie und Internationales)
- Dr. Rolf Oppliger (EFD; ISB/SEC, Informatiksicherheitsbeauftragter Bund)
- Marc-André Ryter (VBS; ASTAB)
- Reto Wollenmann (EDA; Abteilung Sicherheitspolitik; Stellvertretender Sektionschef Rüstungskontrolle, Abrüstung und Nonproliferation)
- Dr. Marco Willisch (GS-EFD – Cyber-Sicherheit; NCS Koordinator, Geschäftsstelle Cyber-Sicherheit)

**Spezifische Auftragsstudien:**

- Etude / Contribution préliminaire au thème Cybersécurité et Politique de sécurité; EPFL; 04-2019.
- Studie Künstliche Intelligenz, technologischer Wandel und nationale und internationale Sicherheitspolitik; CSS ETH Zürich; 04-2019.(nicht publiziert, nicht öffentlich);
  basierend auf:
    - Fischer, S.-C. und Wenger, A. (2019). Ein neutraler Hub für KI-Forschung. *Policy Perspectives* Vol 7/2. Zürich: Center for Security Studies
    - Fischer, S.-C. (2018). Künstliche Intelligenz: Chinas Hightech-Ambitionen. *CSS Analysen zur Sicherheitspolitik* Nr. 220. Zürich: Center for Security Studies
    - Baezner, M., Maduz, L. und Prior, T. (2018). Intelligente Schutzsysteme für die Stadt der Zukunft. CSS Analysen zur Sicherheitspolitik Nr. 235. Zürich: Center for Security Studies

# 1 Übersicht

Der Einsatz und die Weiterentwicklung von KI wird zunehmend auch aufgrund von sicherheitspolitischen Aspekten und Überlegungen beurteilt. Aus sicherheitspolitischer Sicht gibt es drei Themenfelder, in denen KI eine wesentliche und zunehmende Rolle spielt:

- Aussensicherheitspolitik und internationale Gouvernanz
- Streitkräfte und Wandel der Kriegführung
- Nachrichtendienste und innere Sicherheit

In diesem Kontext existieren unterschiedliche Formen von Risiken. Einerseits besteht das Risiko von unerwarteten technischen Unfällen oder systemischen Fehlern, die aufgrund der zunehmenden Vernetzung und Verwundbarkeiten zu gravierenden Störungen oder Schädigungen von **kritischen Systemen** führen können. Andererseits können solche Schädigungen oder Störungen auch durch bewusste Manipulationen herbeigeführt werden. Daneben ist KI zunehmend auch ein Faktor in den internationalen Beziehungen zwischen Staaten und von machtpolitischen Interessen. KI wird von vielen und insbesondere den grösseren Staaten als strategische Ressource wahrgenommen. Im KI-Bereich hat auf globaler Ebene ein technologischer Wettlauf und ein eigentliches Wettrüsten eingesetzt, das die strategische Stabilität und internationale Sicherheit herausfordert und belastet; dazu gehört auch die Problematik der Weiterentwicklung und Verwendung **von Waffensystemen mit zunehmender Autonomie**, die die Kriegführung stark verändern dürften. Autonome Waffensysteme könnten dereinst einen so hohen Grad an Autonomie erreichen, dass sie nicht mehr der notwendigen menschlichen Kontrolle oder Verantwortlichkeit unterstehen, und quasi selbstständig Ziele erkennen, auswählen und angreifen können. Gleichzeitig gibt es Bestrebungen, dank autonomeren Waffensystemen mit Sensoren und Algorithmen präzisere Waffensysteme zu entwickeln, um so Kollateralschäden zu verhindern, um namentlich auch das Kriegsvölkerrecht besser einzuhalten.

Sicherheitspolitisch relevante Fragen wirft auch das **Ungleichgewicht zwischen staatlichen und privaten Akteuren** auf. Die führenden Technologieunternehmen verfügen bezüglich KI-Systemen über einen erheblichen Wissens-, Daten- und Anwendungsvorsprung, der ihnen gegenüber Staaten viel Einfluss verleiht. Ungleichgewichte bestehen aber auch zwischen den Staaten. Auf globaler Ebene sind die Entwicklungen im KI-Bereich, wie auch in konventionellen Bereichen, stark von der machtpolitischen Rivalität zwischen den USA und China geprägt.

Aufgrund der Zunahme von Cyberangriffen gewinnt KI mit «Deep-/Machine-Learning-gestützten Systemen» auch zur Erkennung von **Cyberbedrohungen** sowie zur Prävention und Aufklärung von Cyberangriffen an Bedeutung. Automatisierte Lösungen zur Minimierung von Cyberrisiken können zudem bspw. mittels «Assistenten» für Cybersicherheit-Teams oder automatisiertem Schutz von Netzwerken den Fachkräftemangel im Bereich Cybersicherheit teilweise abfedern.

Sicherheitspolitische Fragen werden auch im Bereich der **Überwachung** aufgeworfen; auch hier eröffnen KI-Technologien neue Möglichkeiten, sowohl was die Menge der überwachten Kommunikationskanäle als auch die Datenmengen (Randdaten und Nutzdaten) betrifft. Das Thema betrifft die grossen internationalen Unternehmen wie auch sicherheitsrelevante Tätigkeiten von Staaten, z.B. im Polizei- und Nachrichtendienstbereich, wo KI-gestützte Systeme zur Prävention von Straftaten sowie zur Strafverfolgung eingesetzt werden.

# 2 Herausforderungen

Allgemein hat KI auf die Cybersicherheit positive wie auch negative Auswirkungen. KI erweitert die Möglichkeit der automatisierten Handlungen von Maschinen und im IT-Bereich. Das kann im negativen Sinne genutzt und missbraucht werden, wie z. B. bei der automatisierten Durchsuchung von Netzwerkstrukturen und -aktivitäten für Spionagezwecke (*Computer Network Exploitation)* oder um *Attacken* effizienter, schneller, präziser durchzuführen. KI-Angebote werden auch genutzt, um effizientere Phishing-Betrügereien, Speicherung von

„Zero-day Exploits[1]“ oder politische Desinformation und Propaganda zu betreiben. Zudem können unausgereifte KI-Methoden selbst zu neuen attraktiven Angriffsflächen führen. Demgegenüber kann sich diese Bündelung von Technologien auch positiv auf Verteidigungsmöglichkeiten auswirken, zumal KI bereits für die frühzeitige Identifizierung von Cybervulnerabilitäten genutzt wird, indem die Integration und Interpretation von grossen Mengen an Sensordaten heute zeitnah erfolgen kann. Neue Angriffsmuster und -vektoren können so besser erkannt und Sicherheitselemente möglichst in Echtzeit (z.B. Malwareschutz, Firewalls) kontrolliert und gesteuert werden.

## 2.1 Aussenpolitik und internationale Gouvernanz

Es geht darum, Einfluss und Hebelfaktoren der KI gegenüber Aussenpolitik und wirtschaftlicher Macht und Ungleichheit zu untersuchen. Folgende Fragestellungen drängen sich auf:

- Inwieweit beeinflussen KI-gestützte Systeme die internationale strategische Stabilität?
- Führen KI-Systeme zu einem Verlust der internationalen Eskalationskontrolle?
- Wie verändert KI das globale Kräftegleichgewicht, z.B. aufgrund der erhöhten Autonomie bei Waffensystemen?
- Welche Herausforderungen bringt KI für die Rüstungskontrolle?

Die internationale Entwicklung und der Einsatz von KI sind in der Aussenpolitik bisher nur punktuell thematisiert worden[2]. Die «*Dual-Use*»-Eigenschaft von KI, also die zivile und nichtzivile Nutzung, ist jedoch eine der Herausforderungen der aussenpolitischen Überlegungen.

Staatliche Akteure sehen KI zunehmend als strategische Ressourcen an. Sie nehmen vermehrt Einfluss auf den Innovationsprozess und die Weiterverbreitung von KI-Technologien, weil sie dies auch als Mittel zur machtpolitischen Positionierung und Einflussnahme erachten. Dies wird verstärkt durch den allgemeinen Trend in der internationalen Sicherheitspolitik, wieder vermehrt und offener auf machtpolitische Mittel und Interessenverfolgung zu setzen. Diese Entwicklung ist für Staaten, die auf eine regelbasierte Ordnung und multilaterale Lösungen setzen, besorgniserregend. Es gibt erste zaghafte Versuche, auch die Thematik KI im multilateralen Rahmen anzugehen. So wird beispielsweise auf UNO-Ebene über ein Verbot autonomer Waffensysteme («*lethal autonomous weapon systems*») debattiert. Grosse Militärmächte wie die USA, China und Russland sind aber gegen ein solches Verbot, so dass diese Verhandlungen schwierig und noch nicht weit fortgeschritten sind.

KI bringt zahlreiche Herausforderungen für die ***Rüstungskontrolle*** mit sich. Diese betreffen unterschiedliche militärische Bereiche, wie beispielsweise Waffensysteme und Munitionstypen. KI könnte zu einem tiefgreifenden militärischen Veränderungsprozess führen, und die bestehenden nationalen und internationalen Normen und Prozesse auf die Probe stellen. Die Herausforderungen manifestieren sich insbesondere bei autonomen Waffensystemen, zumal der Einsatz von KI die Rüstungskontrolle erschwert. Vollautonome Waffensysteme existieren derzeit noch nicht. Allerdings nimmt die Autonomie kontinuierlich zu. Solche Systeme sind militärisch von strategischer Bedeutung, weil sie u.a. präziser eingesetzt werden können, im Vergleich zum Mensch viel schneller entscheiden oder weitreichender wirken können. Übergeordnetes Ziel ist, dass solche hochentwickelten Systeme weniger militärische und zivile Opfer erfordern.

Für die Rüstungskontrolle stellen autonome Waffen eine komplexe Herausforderung dar, weil sie zugleich politische, militärische, rechtliche und ethische Fragen aufwerfen.[3] Eine

---

1 Eine Zero Day Exploit Attack (ZETA) ist ein Angriff, der am selben Tag erfolgt, an dem die hierbei ausgenutzte Schwachstelle in der entsprechenden Software entdeckt wird. In diesem Fall wird die Schwachstelle ausgenutzt, bevor sie vom Softwarehersteller durch einen Fix geschlossen werden kann [Kaspersky.de].

2 vgl. AI Report_OECD_10-2018.pdf; Artificial Intelligence in Society - Phase 2; OECD, Paris 11-2018.

3 vgl. Markus Christen, Thomas Burri, Joseph Chapa, Raphael Salvi, Filippo Santoni de Sio, and John Sullins “An Evaluation Schema for the Ethical Use of Autonomous Robotic Systems in Security Application”, University of Zurich Digital Society Initiative White Paper No. 1, 2017, 89 pp., SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3063617

ethische Befürchtung ist beispielsweise, ob mit autonomen Waffen der Krieg «entmenschlicht» wird und ob der Entscheid zur Tötung einer Maschine delegiert werden darf. Aus rechtlicher Sicht ist z.B. zu klären resp. sicherzustellen, dass bei ihrem Einsatz das Völkerrecht und insbesondere das humanitäre Völkerrecht - und somit die völkerrechtlichen Grundprinzipien wie das Gebot der Verhältnismässigkeit und das Verbot des unterschiedslosen Angriffs – eingehalten werden. Auch wird befürchtet, dass künftig Kriege «günstiger» werden, dass ein KI-Rüstungswettlauf infolge der «Dual-Use»-Problematik kaum kontrollierbar ist oder dass auch nichtstaatliche Akteure solche Waffen erstreben könnten.

Vor dem Hintergrund dieser Herausforderung plädieren einige Staaten für ein Totalverbot von autonomen Waffensystemen als präventive Vorkehrung. Allerdings ist ein solches Totalverbot international hoch umstritten und folglich kein vielversprechender Weg, weil gerade die militärischen Grossmächte ein solches Vorhaben kategorisch ausschliessen. Rüstungskontrollgespräche in der UNO veranschaulichen bestehende Divergenzen betreffend die Gouvernanz und Regelung. Auch wird klar, dass Rüstungskontrollansätze aus der Vergangenheit nicht einfach auf solche neuen Herausforderungen angewendet werden können.

Andere Staaten wiederum verfolgen den Ansatz, an der menschlichen Kontrolle bei der Verwendung von KI-gestützten Waffensystemen festzuhalten. Damit verbunden muss die internationale Staatengemeinschaft den Grad an wünschenswerter, akzeptabler resp. nicht mehr akzeptabler Autonomie definieren. Eine Norm könnte darauf abzielen, Staaten zu verpflichten, die menschliche Kontrolle beizubehalten.

Ein weiterer (Teil-)Ansatz ist, mittels rechtlicher Überprüfungen von Waffensystemen sicherzustellen, dass die Verwendung von neu entwickelten Waffen im Einklang mit dem Völkerrecht ist. Die Überprüfungspflicht entspringt den ***Genfer Konventionen*** (Zusatzprotokoll I, Art. 36) und ist national in der Materialverordnung VBS (SR 514.20) verankert. Allerdings dürfte die Komplexität von Waffenüberprüfungen zunehmen, je mehr KI in ein Waffensystem integriert ist. Es ist international umstritten, ob Waffenüberprüfungen effektiv sind, zumal eine grosse Anzahl Staaten solche Überprüfungen gar nicht durchführen. Zudem ist es unklar, ob man sich auf internationale Standardmethoden und Testprotokolle für die Durchführung der Waffenüberprüfungen einigen könnte.

Es dürfte noch einige Jahre dauern, bis eine gemeinsam getragene Art der Regulierung entwickelt werden kann, die Mehrwert schafft und militärisches und ziviles Potenzial von KI nicht unnötig einschränkt. Allenfalls können in der Zwischenzeit politische Erklärungen und praktische Arbeiten eine Zwischenlösung darstellen, um den genannten Herausforderungen entgegenwirken.

## 2.2 Streitkräfte und Wandel der Kriegführung

Es geht darum, den Einfluss von KI auf das Kriegs- und Konfliktbild[4] zu antizipieren und mögliche Konsequenzen auch für die eigenen Verteidigungsfähigkeiten zu ziehen. Die Fragen lauten:

- Wie beeinflusst KI die militärische Innovation und Weiterentwicklung von Fähigkeiten?
- Inwieweit können KI-Systeme militärisch destabilisierend wirken (aufgrund neuer Möglichkeiten und Geschwindigkeiten in der Operationsführung)?
- Welche Auswirkungen haben KI-gestützte Systeme auf das Gleichgewicht «Offence-Defence»?
- Welche Auswirkungen hat KI auf asymmetrische Konfliktformen?
- Erleichtert KI die Gewaltanwendung von nichtstaatlichen Akteuren?

Über Breite und Tiefe der Auswirkungen von KI auf die künftige Entwicklung der Kriegführung besteht kein Konsens. Es ist jedoch davon auszugehen, dass KI-Systeme zumindest in den Bereichen Informationsauswertung und Führungsunterstützung eine wesentliche Rolle spielen werden. KI-gestützte Systeme könnten hier in Zukunft schneller und effizienter als

[4] vgl. Article Cybersécurité_MPR_02-2018; Military Power Revue "Cybersécurite: la première ligne de défence contre l'impact de l'intelligence articificielle"; 02-2018.

die heutigen Stabsstrukturen grosse Datenmengen analysieren und im Kontext spezifischer taktisch-operativer Situationen für Entscheide der militärischen Führung aufbereiten. Generell dürfte technisch orientiertes militärisches Personal stark an Bedeutung gewinnen.

Das Potenzial KI-gestützter Systeme, Entscheidungen auf der Basis grosser Datenmengen schneller zu treffen, als menschliche Akteure dazu in der Lage sind, kann sich auf die militärische Stabilität auswirken. Ein so genannter «First-Mover Advantage» könnte im Bereich der Einschätzung, der Führung oder der eingesetzten Waffensysteme kompetitiven Druck auf andere Akteure ausüben und damit Rüstungsdynamiken befeuern. Obwohl damit kein zwangsläufiger Anstieg der Kriegswahrscheinlichkeit einhergeht, können Rüstungsdynamiken die Entwicklung des politisch-strategischen Umfelds negativ beeinflussen und im Extremfall direkt in militarisierte Krisen münden.

Inwieweit die Nutzung von KI in Befehls- und Waffensystemen zu einer Verschiebung des Gleichgewichts zwischen Offensive und Defensive führen wird, ist derzeit noch nicht absehbar. Für die Stärkung der Offensive spricht die Möglichkeit, unerwartete Erstschläge auf einer weit umfangreicheren Datengrundlage «aus dem Stand» durchführen zu können. Verbindet sich diese Entwicklung mit dem Einsatz anderer neuer Technologien, die ebenfalls zu einer Beschleunigung der Gefechtsführung führen (z.B. Hyperschallwaffen), kann das die Bedeutung von eigenen Langstreckenwaffen erhöhen, weil sie als einzige Mittel gesehen werden, um feindliche Äquivalente rechtzeitig auszuschalten. Auf der anderen Seite könnten KI-Systeme auch die Aufklärung und Überwachung gegnerischer Kriegsvorbereitungen erleichtern und damit Fähigkeiten stärken, denen eher stabilisierende Eigenschaften zugerechnet werden. Darüber hinaus würde die Fähigkeit, Überlegenheit durch KI mittels präventiver Cyberangriffe zur Lähmung von Führungssystemen oder kritischen Infrastrukturen zu zeigen, als eine neue Form der militärischen Abschreckung dienen.

## 2.3 Nachrichtendienste und innere Sicherheit

Es geht darum, den Einfluss von KI auf Sicherheitsinstrumente des Staates im Inneren, insbesondere im Bereich der nachrichtendienstlichen Tätigkeiten, zu analysieren und deren Chancen und Risiken zu bewerten. Die Fragen lauten:

- Wie wirkt sich KI auf die Tätigkeit staatlicher Nachrichtendienste aus?
- Welche Auswirkungen hat die KI auf das Spannungsfeld von Sicherheit und Freiheit?
- Inwieweit fördert KI die Wirkung von Propaganda und Desinformation?
- Wie wirken sich KI-Systeme auf kriminelle Aktivitäten im Cyberbereich aus (z.B. effizientere Phishing-Betrügereien, effizientere Speicherung von „Zero-day Exploits“)?

Die Kommerzialisierung und Verbreitung von neuen KI-Technologien erhöht deren Verfügbarkeit für staatliche, aber auch für nichtstaatliche Akteure, darunter auch kriminelle Organisationen und terroristische Gruppierungen. Damit verbunden steigt auch das Missbrauchspotenzial von kommerziell verfügbaren und vergleichsweise kostengünstigen KI-Anwendungen. Zu denken ist in diesem Zusammenhang beispielsweise an von KI unterstützte Cyberangriffe oder auf Personen und Organisationen zugeschnittene Propagandainhalte. In einer ferneren Zukunft dürften auch kommerziell verfügbare KI-Anwendungen in Form von selbstfahrenden Fahrzeugen oder autonomen, im Schwarm agierenden Drohnen ein Missbrauchspotenzial für terroristische Akteure aufweisen.

Das Potenzial von KI kann aber auch in der inneren Sicherheit, z.B. bei der Terrorismusbekämpfung, genutzt werden. Auch wenn gegenwärtig in diesem Bereich erst wenige konkrete KI-Systeme verfügbar sind, befinden sich entsprechende Anwendungen in Entwicklung. Dazu gehören Algorithmen zur Identifikation von Personen, durch KI unterstützte Datenanalysewerkzeuge, KI-unterstützte Applikationen im Bereich der Bekämpfung der Terrorismusfinanzierung und KI-unterstützte Methoden, um effizienter gegen Propaganda terroristischer Gruppierungen im Internet vorzugehen. Auf der anderen Seite besteht das Risiko, dass die durch KI erweiterten Möglichkeiten zur Überwachung auch zu illegitimer Kontrolle und repressiven Zwecken eingesetzt werden. Für freiheitlich, rechtstaatlich verfasste Staaten wird sich hier die Frage stellen, bis zu welchem Grad und welchem Zweck der Einsatz solcher

Systeme wünsch- und vertretbar ist. Das latente Spannungsfeld zwischen Sicherheitsüberlegungen und dem Gebot der individuellen Rechte und Freiheiten könnte sich hier weiter akzentuieren.

KI-Systeme und KI-Technologien werden es Nachrichtendiensten ermöglichen, immer grössere Datenmengen zu sammeln und auszuwerten. Die öffentlich verfügbare Datenmenge nimmt aufgrund der Entwicklung des Internet der Dinge, der Verfügbarkeit von sogenannten «Smart Devices» und der umfassenden Onlineaktivitäten von Privatpersonen, Firmen und Verwaltungsstellen exponentiell zu. Dies stellt staatliche Nachrichtendienste vor umfassende Herausforderungen. Die Integration von KI in die nachrichtendienstliche Arbeit kann staatliche Nachrichtendienste effizienter machen, indem sie die Qualität, Verfügbarkeit und Zielorientierung nachrichtendienstlicher Analysen verbessert. Allerdings können verbesserte Überwachungstechnologien in den Händen anderer Nachrichtendienste auch die Durchführung verdeckter Operationen und die Nachrichtengewinnung erschweren.

Mit KI-Systemen können Propaganda und gefälschte Inhalte noch gezielter und mit einer höheren Reichweite und Effektivität verbreitet werden.[5] In den Händen übelwollender staatlicher, halbstaatlicher oder nichtstaatlicher Akteure können sie ein Instrument sein, um die politische Stabilität durch die massenhafte Verbreitung von Desinformation zu untergraben. KI ermöglicht immer realistischere Fälschungen von Foto-, Audio- und Video-Material. Sogenannte «Deep Fakes» können verwendet werden, um falsche Nachrichtenberichte zu generieren, den öffentlichen Diskurs zu beeinflussen, Politiker oder staatliche Institutionen zu erpressen und falsche digitale Identitäten zu erstellen. KI kann aber auch zur Bekämpfung von Propaganda eingesetzt werden. KI-basierte forensische Werkzeuge können Bots blockieren, Fälschungen aufdecken und Desinformation aussortieren. Digitale «Distributed-Ledger»-Technologie und Maschinengeschwindigkeitssensoren können dazu beitragen, Echtzeitinformationen und die Echtheit von Bildern und Videos zu bestätigen und so die Zuverlässigkeit von Informationen zu erhöhen. Wenn es um gefälschte Videos von Personen geht, können KI-gesteuerte Programme aufgrund biometrischer Indikatoren wie Puls und Sprache Fälschungen erkennen. Die Identifizierung von manipulierten Satellitenbildern ist hingegen viel schwieriger.

# 3 Bestehende Aktivitäten

Der Bund hat mit seiner Strategie "Digitale Schweiz" Massnahmen definiert und zum Teil schon implementiert, die es erlauben, die Behörden, die Gesellschaft und die Wirtschaft zu sensibilisieren und auf potenzielle Chancen und Bedrohungen aufmerksam zu machen. Bei der fortschreitenden Digitalisierung soll auch dem Aspekt der Sicherheit gebührend Rechnung getragen werden, vor allem hinsichtlich der kritischen Informations- und Kommunikationstechnologien und -infrastrukturen, die für das Funktionieren von Staat, Wirtschaft und Gesellschaft elementar sind. Deshalb wird schon heute, insbesondere mit den Betreibern kritischer Infrastrukturen, die Zusammenarbeit betrieben und verstärkt.

Das Verfolgen und Analysieren der internationalen sicherheitspolitischen Entwicklungen im Cyberbereich, insbesondere auch mit Bezug zur Entwicklung der Streitkräfte, wird durch das VBS zusammen mit dem EDA wahrgenommen. Internationale Beziehungen und Kooperationen werden diesbezüglich mit verschiedenen Ländern und Organisationen gepflegt und weiterentwickelt.

Im Folgenden werden Instrumente und Organisationen erwähnt, die der Bund etabliert hat und betreibt. Sie decken bereits jetzt sicherheitspolitisch relevante Aspekte von KI ab und erbringen konkrete Leistungen zugunsten der Sicherheit auch im Cyberraum, ebenfalls mit einem Bezug zu KI.

[5] vgl. Wenn Algorithmen für uns entscheiden: die Herausforderungen der künstlichen Intelligenz; Christen et al.; TA-SWISS Publikationsreihe (https://www.ta-swiss.ch/projekte-und-publikationen/informationsgesellschaft/kuenstliche-intelligenz/)

## 3.1 Nationale Strategie zum Schutz der Schweiz vor Cyber-Risiken (NCS)

Die NCS[6] wurde 2012 definiert und wird nun seit 2018 in einer zweiten Phase weiter implementiert. Sie listet eine Reihe von laufenden sowie geplanten Aktivitäten und Projekte auf, die zur Stärkung des Schutzes vor Cyber-Risiken beitragen und auch einen Bezug zu KI haben. KI-Technologien werden heute im Bereich der Cybersicherheit der Bundesverwaltung noch kaum zielgerichtet eingesetzt oder nur sehr punktuell in Form von ausgereiften Produkten (bspw. zur Malware-Bekämpfung). Zu berücksichtigen ist hierbei aber, dass die aktuelle NCS und deren Umsetzungsplanung lediglich für den Zeithorizont bis und mit 2022 ausgelegt ist; bis dahin dürften sich die KI-Technologien weiterentwickeln. Mit der Ernennung eines Delegierten des Bundes für Cybersicherheit und mit der Schaffung des geplanten schweizweit vernetzten ***Cyber-Kompetenzzentrums*** werden durch das EFD im Rahmen der gültigen NCS weitere Voraussetzung geschaffen für eine integrale Handhabung von KI als Querschnittsthema.

Insbesondere mit den Instrumenten MELANI (Melde- und Analysestelle Informationssicherung) und GovCERT (*Governemental Computer Emergency Response Team*) sind beim ISB Fähigkeiten vorhanden, bekannte, aber auch neue Risiken im Cyberbereich zu antizipieren und analysieren, wobei es auch hier bereits eine weitgehende, etablierte Kooperation gibt (insbesondere mit Betreiber kritischer Infrastrukturen).

## 3.2 Nationale Strategie zum Schutz kritischer Infrastrukturen (SKI)

Kritische Infrastrukturen versorgen die Bevölkerung und Wirtschaft mit essenziellen Produkten und Dienstleistungen (z.B. Strom, Treibstoff, Zahlungsverkehr, Transportdienstleistungen, medizinische Versorgung).

Die Einführung von KI bei kritischen Infrastrukturen verspricht nicht nur markante Effizienzgewinne z.B. bei der automatisierten Überwachung von komplexen Anlagen, sondern auch ganz neue Möglichkeiten bei der Erbringung öffentlicher Dienstleistungen (Smart Cities). Neben den genannten Chancen birgt KI jedoch auch erhebliche Risiken.

Die Risiken können durch eine aufgrund von öffentlichem und/oder wirtschaftlichem Druck vorschnellen Implementierung von KI mit einseitigem Fokus auf die Chancen und Ausblendung der Gefahren verschärft werden. Ein zentrales Risiko liegt in der deutlich zunehmenden Komplexität und Undurchschaubarkeit von KI-Systemen. Dieses Risiko wird durch die zunehmend komplexen Versorgungsketten insb. im Cyberbereich weiter erhöht ((Cyber) Supply Chain Security). Ein weiteres Risiko besteht in der drohenden Fähigkeitserosion bei Mitarbeitenden innerhalb der kritischen Infrastrukturen. Da KI u.a. dazu verwendet werden soll, Systeme automatisiert und nach bestimmten Algorithmen selbständig zu überwachen sowie zu beeinflussen, verändert sich die Rolle des Menschen in diesen Systemen zunehmend weg von einem aktiven hin zu einem passiven Systemelement. Risiken frühzeitig zu erkennen, aktiv zu behandeln und bei Störungen kompetent reagieren zu können, wird jedoch auch in Zukunft eine Kernaufgabe des Betriebspersonals sein, erfordert jedoch entsprechende Kompetenzen, ein umfassendes Systemverständnis wie auch ein adäquates Situationsbewusstsein.

Auch im Katastrophenfall bringt KI ein erhebliches Potenzial mit sich. Die Zurverfügungstellung von zeitaktuellen, intelligent aufbereiteten Daten können Entscheidungsfindungsprozesse von Führungsorgangen bei der Ereignisbewältigung massgeblich unterstützen. Dies jedoch nur dann, wenn die Versteh- und Handhabbarkeit der aufbereiteten Daten den kognitiven Fähigkeiten von Menschen entspricht. Dies ist eine Frage der Benutzerfreundlichkeit, die sich bei der Entwicklung der Systeme stellt und praktische Tests erfordert.

Durch die Einführung von KI nimmt die Geschwindigkeit und Dynamik von Interaktionen zwischen technischen Systemen innerhalb einer Organisation wie auch in Verbindung mit Systemen ausserhalb der Organisation markant zu. Entwicklungen und dynamische Prozesse,

[6] vgl. Nationale Strategie zum Schutz der Schweiz vor Cyber-Risiken (NCS) - 27.6.2012 und Nationale Strategie zum Schutz der Schweiz vor Cyber-Risiken (NCS) 2018-2022 - 18.4.2018

die teilweise kaum oder gar nicht vorhersehbar und verstehbar sind, können ein erhebliches Risiko für das normale Funktionieren von kritischen Infrastrukturen darstellen. Fehler und Sicherheitslücken, die in den Quellcodes von KI-Systemen möglicherweise einprogrammiert sind, können im operativen Betrieb kaum entdeckt werden und bilden das Einfallstor für unerwünschte externe Beeinflussungen. Durch die zunehmende Vernetzung wird dieses Risiko zusätzlich verstärkt, da nicht davon auszugehen ist, dass jede in einer Versorgungskette beteiligte Organisation über gleich hohe Cybersicherheit-Fähigkeiten verfügt. Entsprechend schlechter aufgestellte Organisationen bilden Schwachpunkte und somit Einfallstore innerhalb dieser (Cyber-)Versorgungsketten. Dies kann durch die starke Vernetzung und Automatisierung durch KI-Anwendungen negative Auswirkungen auf das Funktionieren von kritischen Infrastrukturen haben. Dieses Risiko steigt insbesondere dann, wenn Betreiber von kritischen Infrastrukturen Aufgaben an Dritte auslagern und kein hinreichendes Cybersicherheit-Konzept zur systematischen Behandlung von Verwundbarkeiten vorliegt.

Mit der SKI wurden Massnahmen für einen besseren Schutz von kritischen Infrastrukturen und der Versorgung mit entsprechenden Dienstleistungen für den Zeitraum 2018–2022 verabschiedet (inkl. kritische Informations- und Kommunikationsinfrastrukturen). Das BABS verfügt mit der SKI über ein Instrument, mit dem auch neue, zum Beispiel durch KI verursachte Risiken für kritische Dienstleistungen, aber auch Chancen für besseren Schutz erfasst werden können.

Versorgungsketten (inkl. Cyber Supply Chain) müssen vor allem im Krisenfall im Verbund gesichert werden. Dies verlangt vergleichbar hohe Cybersicherheit-Fähigkeiten bei allen Partnern, was in der Realität schwierig sicherzustellen ist, auch weil kritische Infrastrukturen oft von Dritten betrieben werden.

Aus KI-Sicht ist vor allem die Datenpflege relevant, weil spezifische aufbereitete Daten der Schlüssel für einen zielgerichteten und effizienten Einsatz von KI zur Unterstützung von Führungsorganen im Not-/Katastrophenfall sind. Interoperabilität wird immer mehr an Bedeutung gewinnen. Mit dem ***Sicherheitsverbund Schweiz (SVS)*** als Bindeglied zwischen dem Bund und den kantonalen Behörden sollte die Interoperabilität auch verstärkt unter dem Blickwinkel der KI verfolgt werden. Als Beispiel dient die Harmonisierung der Schweizer Polizeiinformatik (HPi).

## 3.3 Aktionsplan Cyberdefence (APCD)[7]

Mit der Definition und der Umsetzung des APCD sorgt das VBS seit 2016 vor allem für den Schutz seiner eigenen Netze und Systeme. Als für die Sicherheitspolitik federführend zuständiges Departement muss es in der Lage sein, die in Anzahl, Intensität und Komplexität zunehmenden Formen von Cyberbedrohungen zu bewältigen, sowohl im Alltag als auch im Fall einer Krise oder eines Konflikts. Weiter gilt es, Cyberaspekte des Nachrichtendienstgesetzes (NDG) und des Militärgesetzes (MG) konkret umzusetzen und die dafür nötigen Fähigkeiten aufzubauen. Das VBS ist im Cyber- und damit im KI-Bereich als gesamtheitliches System zu verstehen. Dieses umfasst den NDB und das BABS für das zivile Umfeld, die Armee für die militärischen und technischen Aufgaben und armasuisse für die Beschaffung und Forschung. Das VBS kann mit Mitteln der Armee Betreiber kritischer Infrastrukturen und zivile Behörden, bei Bedarf auch in Folge von Cyberangriffen, subsidiär unterstützen. Eine Cyberdefence Expertengruppe (Wissensträger aus der Hochschulen, Verwaltung und Wirtschaft) kann zur strategischen Beratung beitragen.

## 3.4 Cyberdefence-Campus (CYD-Campus)

Für die Antizipation und die Zusammenarbeit mit den Hochschulen und Forschungszentren, auch international, wurde der im APCD definierte CYD-Campus im VBS bei armasuisse

[7] vgl. Aktionsplan Cyberdefence VBS (APCD) 09.11.2017 und Aktionsplan Cyberdefence VBS (APCD) Revision 17.12.2018.

W+T[8] zu Beginn 2019 gegründet. Der CYD-Campus definiert sich als Teilorganisation innerhalb eines Netzwerks, welche die Interessen der Bundesbehörden (im Fokus steht das VBS) in den Bereichen Cybersicherheit und Cyberdefence mit Leistungen der Akademie (d.h. Hochschulen) und der Wirtschaft (d.h. Industrie) verbindet bzw. vermittelt. Der CYD-Campus ist eine Plattform zur Früherkennung und zum Monitoring von neuen Technologien und hat unter anderem die Aufgabe, Konsequenzen der rasanten Entwicklungen, darunter auch KI-Entwicklungen, zu antizipieren und abzuleiten für:

- die operationellen Fähigkeiten der Armee;
- das Marktumfeld im Sicherheitsbereich;
- das Verhalten von Individuen und Organisationen.

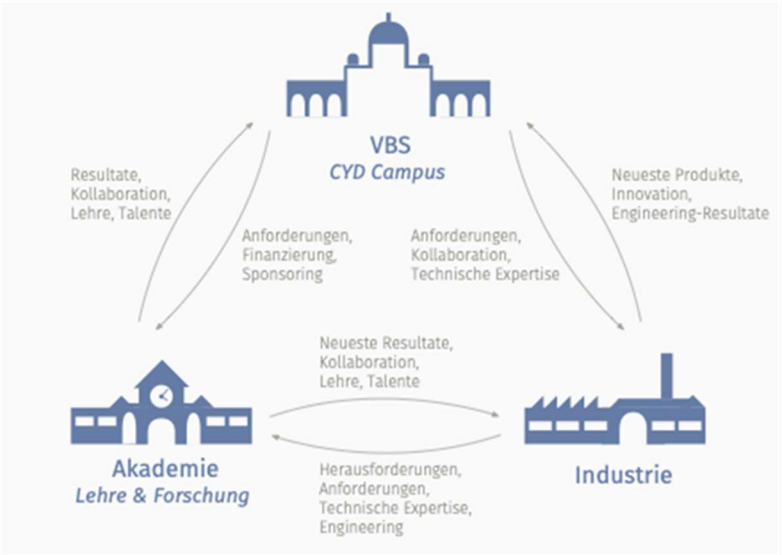


*Abbildung 3: Der Cyberdefence-Campus, ein schweizweites Netzwerk für die Koordination und Stärkung von Forschungs- und Bildungsaktivitäten im Cyberbereich, unter Leitung der armasuisse W+T.*

Die Zusammenarbeit mit den beiden ETH Zürich[9] und Lausanne [10]erlaubt es, auch von deren Knowhow und Initiativen zu profitieren. Als Beispiel kann der *Swiss Data Science Center* (SCSC) erwähnt werden. Ziel des SCSC ist es, Experten der *Data Science,* als Teilbranche der KI, mit Fachspezialisten der Anwendungsdomänen in Kontakt zu setzen; denn nur die Kombination dieser Gebiete ermöglicht es, KI mit nutzbringend einzusetzen. Gemeinsame angewandte Forschungsprojekte werden zurzeit definiert, die KI integrieren, wie z.B. Lösungen für die Erfassung der integralen Cybersicherheitslage oder Testumgebungen für das Prüfen und Zertifizieren von Algorithmen (*Digital Science*). Die beiden ETH bieten auch einen gemeinsamen *Master in Cybersichersicherheit* an, der auch KI relevante wissenschaftliche Bereiche abdeckt. Der Ansatz des *Computational Thinking* spielt für KI eine bedeutende Rolle.

Mit Bezug auf KI verfolgt der CYD-Campus folgende Reihe von Aktivitäten:

- **Antizipation von Cybertechnologien:** Kartographie von Cybertechnologien und -akteuren mit Hilfe von Technologie- und Markt-Monitoring (TMM). Der Bezug zu KI ist hier prominent.
- **Früherkennung** technologischer Entwicklungen aus der KI, die im Rahmen der Cybersicherheit von Bedeutung sind.
- **Führung und Koordination** von Forschungsaktivitäten im Rahmen der Cybersicherheit. Der Bezug zu KI ist wiederum in vielfältiger Weise gegeben.

8 armasuisse Wissenschaft und Technologie.
9 vgl. Studie KI und Sicherheitspolitik_CSS_04-2019; Künstliche Intelligenz, technologischer Wandel und nationale und internationale Sicherheitspolitik; CSS ETH Zürich; 04-2019.
10 vgl. Cybersécurité et Politique de sécurité_EPFL_04-2019.

- **Planung** von technischen und wissenschaftlichen Fach-Konferenzen. Eine KI-Fachkonferenz ist für Mitte September 2019 in Lausanne geplant.
- Aufbereitung, **Publikation** und Verbreitung von Erkenntnissen, die aus den Aktivitäten der Antizipations-Plattform stammen.
- Ableiten von (**disruptiven) Konsequenzen** der technologischen Entwicklungen, insbesondere auch KI-Entwicklungen, für:
    - die operationellen Fähigkeiten der Armee
    - das Marktumfeld im Sicherheitsbereich
    - das Verhalten von Individuen, Organisationen und Staaten
    - gesellschaftliche und politische Entwicklungen
- **Ausbau und Weiterentwicklung der Kompetenz-Netzwerke**:
    - Förderung internationaler Kooperationen (u.a. mit CCDCoE, EDA, NATO/PfP), auch im Bereich KI.
    - Organisation der CYD-Campus-Seminare, die jeden zweiten Montag in Thun zur Lunchzeit stattfinden und vielfältige Beiträge aus der KI von Vertretern der Industrie, von Universitäten und Cyberexperten bieten.
    - Zugang zur Forschungsplattform, bestehend aus dem Data Science Lab und einem GPU-Cluster-System (für den Bereich Machine Learning) zur Entwicklung und Erprobung gemeinsam erarbeiteter Demonstratoren. (Geplant ist die Etablierung eines Cyberlabs mit unterschiedlichen Vertraulichkeitsstufen beim Zugang, der Datenhaltung und Informationsverarbeitung.
    - Förderung agiler Entwicklungs- und Implementierungs-methoden im militärischen Umfeld, die im Rahmen von KI-Anwendungen von großer Bedeutung sind, inkl. Prüfung alternativer Beschaffungsprozesse.
    - Leitung von Bachelor-/Master-Arbeiten oder Doktoraten in Zusammenarbeit mit unterschiedlichen Universitäten und zu Themen, die im Rahmen der Cybersicherheit oder auch Data Science relevant sind.
    - Bereitstellung von Ausbildungsmodulen zu Cybersicherheit und Data Science für unterschiedliche VBS-Stellen. KI besitzt in diesen Lehrgängen einen großen Stellenwert.
    - Aufbau und Bereitstellung attraktiver Ausbildungs- und Stellenangebote für Miliz-, Berufs- und Zivil-Funktionen innerhalb des VBS.
    - Frühe Identifikation von Talenten zur Rekrutierung, z.B. durch Organisation von Hackathons oder Wettbewerben.
- **Entwicklung und Applikation**
    - Aktive Führung von Forschungsprojekten in zwei Themenbereichen. Einbezug in deren Umsetzung von sowohl Universitäten wie auch Vertreter der Industrie.
      Im ersten Themenbereich *Cybersicherheit* setzen die aktuellen Forschungsprojekte mehrheitlich KI-Technologien ein.
      Im zweiten Themenbereich *Data Science* beschäftigt sich eine Reihe von Forschungsprojekten dediziert mit KI-Themen, wie beispielsweise der Bild-Klassifizierung mittels Deep Learning, oder mit prädiktiven oder kausalen Modellen, welche KI-Methoden verwenden.
    - Die Resultate aus den Forschungsaktivitäten kommen in Form von Publikationen oder Referaten einer breiten Cyberdefence Community zugute.
    - In Form von Demonstratoren dienen sie unterschiedlichen Stakeholdern aus dem VBS als Proof of Concept oder als Prototyp zwecks einem Know-how- oder Technologie-Transfer für operative Systeme.

## 3.5 Cyberspezialisten in der Armee (Cyberlehrgang)

Dank des Milizsystems verfügt die Armee mit dem 2018 geschaffenen ***Cyberlehrgang*** über ein zusätzliches Mittel, um Hochschulabsolventen mit aktuell erworbenen Kompetenzen zu rekrutieren und in die Cybereinheiten der Führungsunterstützungsbasis, inkl. elektronischer Operationen, zu integrieren. Dies soll der Armee auch dabei helfen, verstärkt KI-Aspekte und Entwicklungen (inkl. «*Dual-Use*») einzubeziehen und zu verfolgen.

Die Verfolgung der wirtschaftlichen Innovation im Bereich Cybersicherheit wird vom Bund mit Beiträgen des VBS an Fachevents gepflegt, wie zum Beispiel mit den *Swiss Cyber Security Days 2019*.

### 3.6 Internationale Zusammenarbeit

Das EDA hat 2018 das Büro für Cyberaussen- und -Sicherheitspolitik etabliert, um die cyberspezifischen aussen- und sicherheitspolitischen Interessen der Schweiz gegenüber ausländischen Partnern und internationalen Organisationen kohärent und konsistent zu wahren.

Auf technisch-operativer Ebene arbeitet das ***govCERT des ISB*** mit internationalen Partner zusammen, um Informationen zur Vorfallbewältigung auszutauschen. Das govCERT ist Mitglied von europäischen und globalen technischen Netzwerken, wie z.B. das Europäische GovCERT (EGC). Diese Zusammenarbeit erlaubt es den technischen Organisationseinheiten, technologische Entwicklungen, inkl. KI und die damit einhergehenden Bedrohungen, zu identifizieren.

Im *Cyberdefence*-Bereich kooperiert die Armee bilateral insbesondere mit Nachbarländern wie Frankreich, Deutschland und Österreich. KI entwickelt sich zunehmend als fixes Thema dieses Austausches. Abkommen für militärische Ausbildung oder für den Informationsaustausch werden laufend mit Elementen des Cyberbereichs erweitert.

Die multilaterale Kooperation wird auch gezielt angestrebt und verstärkt. Mit der Beteiligung der Schweiz am **C*ooperative Cyber Defence Center of Excellence* (CCDCOE)** in Tallinn/Estland (ab Mai 2019) soll die Zusammenarbeit mit Nato-Staaten und weiteren, vor allem europäischen Partnern verstärkt werden. Das CCDCOE bietet im Bereich Cyberdefence und Cybersicherheit Ausbildungsprogramme und Weiterbildungsmöglichkeiten an. Die Aufgaben des Zentrums umfassen zudem die Planung und Durchführung von Übungen, die Erstellung von Konzepten sowie die Publikation von Forschungsergebnissen und juristischen Handbüchern. Das Zentrum befasst sich zunehmend auch mit dem Thema KI. An das CCDCOE delegierte Fachleute aus der Schweiz werden sich auch gezielt mit KI-relevanten Aspekten wissenschaftlich-technisch und rechtlich befassen können. armasuisse W+T beteiligt sich zudem an gewissen technologischen Forschungsgruppen der europäischen Verteidigungsagentur (EVA), die sich mit Cybersicherheit und KI befassen.

### 3.7 Das Genfer Zentrum für Sicherheitspolitik (GCSP)

Das Genfer Zentrum für Sicherheitspolitik ist eines der drei Kompetenzzentren für Aussen-, Sicherheits- und Friedenspolitik. Es bietet praxisorientierte Lehrgänge und Weiterbildungsprogramme an. KI ist auch Teil der vom GCSP angebotenen Inhalte.

## 4 Bewertung und Handlungsbedarf

Der Bund verfügt bereits über bestimmte Strategien und eingeführte Instrumente, um auf die Herausforderungen durch das Aufkommen von KI-Technologien zu reagieren, welche heute reif genug sind, um Auswirkungen auf die Sicherheits- und Aussenpolitik sowie die Cybersicherheit zu haben. Die Identifizierung von Akteuren mithilfe KI in den betreffenden Bereichen steckt jedoch noch in den Kinderschuhen. Es ist festzulegen, wie die transversale Dimension der KI zu verstehen ist.

### 4.1 Aussensicherheitspolitische Auswirkungen

Die sicherheitspolitische Bedeutung von KI sollte in politischen Grundlagendokumenten künftig stärker berücksichtigt werden. Das Thema wird, zusammen mit der gesamten Cyberproblematik, im nächsten Bericht des Bundesrates über die Sicherheitspolitik der Schweiz behandelt. Dieser ist für den Zeitraum 2020-21 vorgesehen. Es wird dort darum gehen, die sicherheitspolitische Bedeutung des Themas zu erläutern und allfällige Konsequenzen für

die Sicherheitspolitik der Schweiz abzuleiten. Der Einsatz von KI wirft Fragen zur strategischen, internationalen Stabilität auf. Nebst den zahlreichen Vorteilen ist das Missbrauchs- und Manipulationspotenzial gross. Die Herausforderungen sind vielseitig: erschwerte Rüstungskontrolle, zunehmende Autonomie von Waffensystemen, Ungleichgewicht zwischen staatlichen und privaten Akteuren und allgemeines Wettrüsten. KI wird vermehrt als Instrument zur Machtprojektion und politische Einflussnahme genutzt.

Es soll durch das EDA die aussenpolitischen Implikationen, die mit dem Einsatz von KI-gestützten Systemen einhergehen geprüft werden. Dabei ergeben sich folgende aussensicherheitspolitische Fragestellungen:

- Regulierung: Braucht es neue Regulierungsmodelle? KI umfasst politische, rechtliche und ethische Fragen, die internationale Regulierungen verlangen. Aber: Regulierungsprozesses sind statisch und langwierig. Sie können den rasanten Veränderungen im Bereich KI kaum gerecht werden.
- Cybersicherheit: Wie wird AI zur Durchführung komplexer Angriffe genutzt und wie wird dadurch die Bedrohungslage verschärft? Wie wirkt umgekehrt AI als Chance, Attacken besser prognostizieren, verhindern und bekämpfen zu können?
- Normen: Welchen Einfluss auf die Aussensicherheitspolitik und das humanitäre Völkerrecht hat KI? Wer trägt die Verantwortung für die Handlungen mit KI? Wie kann verhindert werden, dass autonome Waffen in terroristische Hände fallen? Welche sind dabei die militärischen, rechtlichen und ethischen Herausforderungen?
- Menschenrechte: Werden Grundrechte (z.B. neues KI-gestützte System in China zur Überwachung der Bürger in nahezu allen Lebensbereichen) durch soziale Kontrolle mit Zensur und Überwachung der Bevölkerung zu sehr kleinen Kosten verletzt?

Die Entwicklung des GCSP als Plattform für den Austausch, auch zur KI mit ihren Auswirkungen auf die Sicherheitspolitik und die internationale Cybersicherheit, muss gefördert werden. Eine klare Differenzierung soll dem Label Schweiz ermöglichen, als Referenz für andere ausländische Initiativen zu dienen. Dies kann im Rechtsgebiet der Sicherheitstechnologien oder bei der Zertifizierung von Massnahmen im Zusammenhang mit der Cyber-Supply-Chain geschehen. Aufgrund ihrer Positionierung und Anerkennung als glaubwürdiger Staat mit hohen technologischen Standards stehen für die Schweiz Chancen im Rahmen von KI eine Wissenschaftsdiplomatie[11] zu entwickeln und zu fördern. Die Verbindung von Wissenschaft mit Diplomatie könnte ein vielversprechendes Instrument sein.

Es soll überlegt und strategisch entschieden werden, ob die Schweiz im KI Bereich spezifische Policy aufbauen soll, ob Policy-Formulierung notwendig und erwünscht sind und ob Aufholbedarf besteht.

## 4.2 Bedrohungsformen und Doktrin

Die Beobachtung der Entwicklung der Mittel und Einsatzdoktrinen ausländischer Streitkräfte und Sicherheitskräfte, einschliesslich des nachrichtendienstlichen Aspekts, muss eine KI-Komponente beinhalten. Das diplomatische Netzwerk und die Verteidigungsattachés müssen informiert werden und diesen besonderen nachrichtendienstlichen Bedürfnissen Rechnung tragen. Insbesondere müssen folgende Fragen beantwortet werden können:

- Wie werden Waffen an der Schnittstelle zwischen Robotik und KI die Kriegführung verändern?
- Wie werden Aufgaben von Soldaten zukünftig von KI übernommen (Bild- und Spracherkennung, sowie das Planen, Ausführen und Optimieren von Operationen).

Auf nationaler Ebene muss die Beschaffung von militärischen Systemen die Möglichkeiten der Aufrüstung und Nutzung von KI berücksichtigen. Es ist ein Mapping der Cyberversorgungs- und -beschaffungsketten von militärischen und nachrichtendienstlichen Systemen zu erstellen.

[11] vgl. Policy Perspectives PP7_CSS_02-2019-D; Ein neutraler Hub für KI-Forschung, ETHZ / CSS; 03-2019.

Erhöhung der Sicherheit durch bereits ausgereifte verfügbare IT-Sicherheitsprodukte; bspw. bei der präzisen Spam-Bekämpfung, Härtung von Systemen, zur Aufspürung von Anomalien in Anwendungen/Systemen/Netzen des VBS und dessen Partner (kritische Infrastrukturen) die auf Schwachstellen oder einen Angriff hindeuten. Entwicklung von spezifischen Cyberfähigkeiten (für Schutzmassnahmen oder zwecks Unterstützung von Operationen durch den Cyberraum) durch Einsatz von KI-Technologien.

## 4.3 Fähigkeiten und Kapazitäten

Die Sicherheitsinstrumente der Schweiz müssen durch KI wo notwendig und möglich gestärkt werden. Die enge Zusammenarbeit mit der Industrie und der Wirtschaft ist ein Schlüsselfaktor für die erfolgreiche Wahrnehmung des KI-Potentials. Deshalb geht es darum, agiler in der Erkenntnis und Formulierung der Bedürfnisse sowie rascher in der Beschaffung und Einführung von Lösungen zu werden (*Translationale Forschung*).

Es geht darum, alle Dimensionen der von der KI betroffenen Cyberwelt, nämlich die Cybersicherheit, die Cyberverteidigung und die Cyberkriminalität, integrieren zu können. Mit dem APCD verfügt das VBS über eine Strategie und Instrumente, die ihm die Flexibilität geben, sich den Herausforderungen der KI zu stellen. Da die Ressourcen sehr begrenzt oder ungleich verteilt sind, sind sie durch bereichsübergreifende KI-Projekte zusammenzuführen.

Gemäss der NCS, respektive dem APCD, kann ein systematischer Ausbau der IT-Sicherheit und Cyberdefence Fähigkeiten mit KI-Technologien ermöglicht werden. Dabei können bereits heute im KI-Bereich bestehenden und etablierten Kompetenzen (Netzwerke von Forschern und Fachkräften, Innovationszentren, Förderinstrumente usw.) der Schweiz dienen.

Mit dem CYD-Campus, respektive mit dem zukünftigen Cyber-Kompetenzzentrum, soll die Schweiz über ein KI-Prüfstand verfügen, der es erlaubt KI-Algorithmen in einer sicheren Umgebung (Lab/test-bed) vorerst zu trainieren, stetig zu überprüfen und weiterzuentwickeln. «Integre» und ausreichende Mengen an Trainingsdaten werden somit essentiell für den KI-Einsatz. Das Sammeln, Strukturieren und Schützen dieser Daten soll zu einer wichtigen und zentralen Aufgabe werden.

In den operativen technischen Stellen, die für die Netzüberwachung des Bundes respektive der Armee verantwortlich sind, soll überprüft werden, wie KI als Analyseunterstützung bei komplexen Aufgaben mit grossen Datenmengen; bspw. bei der Attribution von Angriffen (Clusteranalyse), Malware-Analyse, Security-Testing usw. progressive eingeführt und weiterentwickelt werden.

## 4.4 Antizipation durch Zusammenarbeit, Forschung und Prüfstände

Auf der Grundlage ihrer strategischen Ziele in der KI ist es für die Schweiz wichtig, die zu bevorzugenden Partnerschaften zu identifizieren. Sie verfügt bereits über verschiedene Instrumente zur Unterstützung und zum Ausbau dieser Partnerschaften. Mit der Umsetzung der NCS 2 muss das Kompetenzzentrum Cybersicherheit zusammen mit dem EDA und dem VBS diese Tätigkeiten koordinieren, um den Bund national und international richtig zu positionieren.

Es ist notwendig, die internationale Kooperation auf die Entwicklung eines nachhaltigen und tragfähigen Gouvernanz-Rahmens auszurichten, um die mittel- und langfristige Nutzung von KI nach schweizerischen Grundprinzipien, wie die der Transparenz, Inklusivität und Legitimität zu gestalten.

Die Schweiz hat in der Regel bereits europäische Nachbarländer als Partner. Es wäre jedoch notwendig, den Ausbau des Netzwerks durch andere Partnerschaften, beispielsweise mit technologisch dynamischen Ländern in Asien oder im Nahen Osten, in Betracht zu ziehen, um über differenziertere Sichtweisen zu verfügen.

Die Entwicklung von Forschungs- und Austauschaktivitäten des CYD-Campus mit dem Ausland sollte es ermöglichen, dieses Instrument in Zusammenarbeit mit den spezialisierten Instituten der Hochschulen als Mittel für die internationale Zusammenarbeit und technische Beobachtung im Bereich der sicheren KI zu stärken.

- Das VBS und das WBF sollen im Rahmen ihrer Kompetenzen die Zusammenarbeit mit leitenden Bildungs- und Forschungsinstituten stärken und die Agilität ihrer Antizipation durch Forschungsprojekte (Förderung einer KI translationale Forschung) steigern;
- Das integrale Cyberlagebild soll Entwicklungen im KI Bereich berücksichtigen;
- Durch gezielte Teilnahmen an internationalen Gremien und Forschungsinitiativen kann sich die Schweiz als tragender Partner im KI Bereich positionieren;
- Die Schweiz könnte über das durch das EFD künftig getriebenen Kompetenznetzwerk über einen nationalen KI-Prüfstand verfügen.