VDW Berlin 14.11.2025

Künstliche Intelligenz

Workshop F

Militärische Vertrauensbildung

Moderator: Wolfgang Richter
Oberst a.D.
Associate Fellow GCSP

Künstliche Intelligenz

Militärische Vertrauensbildung

Nutzen von KI-Anwendungen

- Vielfältige Assistenzfunktionen:
- Verarbeitung großer Datenmengen
- m. geringem Zeit-/Personalaufwand
- Eignung: aufwendige Routinearbeiten
- Komplexe Ablaufpläne m. Variablen
- „*Autonome*“ Gestaltung v. Testreihen
- „*Autonome*“ Auswahlvorschläge

Risiken von KI-Anwendungen

- „Vereinfachung“, Beschleunigung >>> Abbau v. Ergebniskontrolle/ Verständnis & Nachvollziehbarkeit komplexer Abläufe
- **Fehlbewertungen** v.a. in sensiblen Bereichen wie *Rechtsentscheidungen Politisch-kulturell-ethischen Analysen*
- **Abhängigkeit von Datengrundlagen**

- Zeitkritische Entscheidungshilfen
- Automatisierte Entscheidungen

- **Verlust menschlicher Kontrolle**
- **Fehler m. schwerwiegenden Folgen**

Künstliche Intelligenz

Militärische Vertrauensbildung

Nutzen militär. KI-Anwendungen

- Rasche Verarbeitung gr. InfoMengen für zeitkritische Entscheidungen
  - *ND-Info, Sat-/Dro-/Elo-/GefAufkl*
- *„Autonome“* Ziellisten, Priorisierung
- *„Autonome“* Auswahlvorschläge für Komplexe OpPläne mit x Variablen
- *„Autonome“* Erstellung v. OpPlänen ??

Risiken militär. KI-Anwendungen

- Weniger Kontrolle automat. Eingriffe in komplexe Ops bei rascher Lageänderung
- **Fehlbewertungen** >>> Gefährdung
  - der ***Auftragserfüllung*** und des ***HVR***
  - der **Verantwortung** der Kommandeure
  - **ethisch-polit.** Entscheidungsfindung
- **strategisch-nukleare Destabilisierung**

- Zeitkritische ***Entscheidungshilfen*** bis
- ***automatisierte Handlungsabläufe***
  - *Raketenabwehr – Aufkl-Schlagkomplex*
  - ***Lethal AWS – „Launch on Warning“ ??***

- **Verlust menschlicher Kontrolle =**
- **Verlust verantwortbaren Handelns**
- **Eskalation mit existentiellen Risiken**

Künstliche Intelligenz

Militärische Vertrauensbildung

Eskalationsvermeidung im Kontext

Frieden u. Spannungszeiten:

Routine-Ops, Patrouillenflüge, Manöver

- **Keine automatisierten Reaktionen!**
  - z.B. bei Laserzielerfassung
  - bei Grenzverletzungen (Flz/Dro)
  - aber: *Schutz gegen RakAngriffe?!*
- **Mil-to-Mil Kommunikation d. op HQs**

- **InfoAustausch: Neue Waffensysteme (Vorstellung?) und Einsatzdoktrinen**
- **Vorführungen der Funktionsweise**

Instrumente der Risikoreduzierung

- **Incident at Sea-Agreements (IncSea)**
  Verantwortbares Handeln der Schiffs- und Flugzeugbesatzungen
- **Ergänzung? – Multilateralisierung?**

- **Wiener Dokument:** Risikoreduzierung Manöverankündigung - **Vorstellung neuer Waffensysteme - MilDoktrinSem**
  keine Implementierung, *kein FSK-Dialog*
- **Wiederaufnahme des Dialogs ?**
- **Ergänzung des Dokuments ?**

Künstliche Intelligenz

Militärische Vertrauensbildung

Eskalationsvermeidung im Kontext

Frieden u. Spannungszeiten:

Cyberraum

- **Keine automatisierten Reaktionen!**
  - z.B. kein automat. *Hack-back*
  - keine Angriffe auf KRITIS
  - **keine Angriffe auf (Nuk) EW/IC³**
- **Souveränität, Aggression-Attribution?**
- **Geltung Humanitäres Völkerrecht**
- **Schutz Internet als „global common“**
- **Verhaltenskodex: Risikoreduzierung**
- **Kooperation gegen kriminelle Akteure**

Instrumente der Risikoreduzierung

- **OSCE Cyber CSBM („voluntary“)**
  - Internet als gemeinsam. Besitzstand
  - Info-Austausch zu gem. Bedrohung
  - Keine Angriffe auf KRITIS
  - gem. Verwundbarkeitsanalysen
  - Kooperation geg. kriminelle Akteure
  - OSCE als Plattform für Dialog
  - Austausch national. CyberStrategien
  - OSCE Communication Network
- **Ergänzung? – Multilateralisierung?**
- **Von VSBM zur Rechtswirksamkeit?**

Künstliche Intelligenz

Militärische Vertrauensbildung

Eskalationsvermeidung im Kontext

Bewaffnete Konflikte

Taktisch-operative Ebene

- Geltung Humanitäres Völkerrecht
- Schutz Zivilbevölkerung, ziv Objekte
- Verhältnismäßigkeit mil. Operationen
- Verantwortung der Kommandeure
- Keine automatisierten OpPläne
- **Keine Zielauswahl und Feuerbefehle ohne menschliche Kontrolle**

Instrumente der Risikoreduzierung

- **Genfer Konventionen, ZP I, CCW**
  - Testpflicht vor Einführung neuer Waffen
  - Müssen HVR-Anforderungen genügen
  - Übungen, Ausbildung der Streitkräfte
- **Zusatzerklärung zum KI-Einsatz?**
- **Beschränkung von LAWS gescheitert !**

- **OSZE Verhaltenskodex**
  Bindung an das Human. Völkerrecht
  Ausbildung der Sicherheitskräfte
- **Ergänzung? – Multilateralisierung?**
- **Zusatzerklärung zum KI-Einsatz?**

Künstliche Intelligenz

Militärische Vertrauensbildung

Eskalationsvermeidung im Kontext

Nuklearstrategische Stabilität

- Zweitschlagfähigkeit schreckt vor Erstschlag ab (MAD), sofern
- genügend Systeme überleben können,
- RakAbwehr / LRF nicht destabilisieren,
- (SAT gest) EW / ISRC³ funktionieren.

- **Keine automatis. Startbefehle für NW („*LAW*“) ohne menschliche Kontrolle**
- **Keine Angriffe > (SAT gest) EW / ISRC³**
- **Wahrung strat. Stabilität: *China*, *LRF* ?**
- **Begrenzung RakAbw - *Golden Dome*?**

Instrumente der Risikoreduzierung

- **New START bis 5.2.2026 – Nachfolge ?**
  - Strategisches Dreieck mit China?
  - strat. RakAbw / **LRF** / **nucSLCM** / ASAT ?

- Zukunft **CTBT Teststopp** / **NVV Art. VI** ?

- **P 5 Erklärung vom 5. Jan 2022**
- **No „*Launch on Warning*“ - Erklärung ?**

- **Outer Space Treaty: Keine NW / WMD**
- **No „*Attacks on SatInfra*“ - Erklärung ?**

- **HCOC:** Freiw. Notifikation Raketenstarts
- **Zusatzerklärung zum KI-Einsatz?**

Künstliche Intelligenz

Militärische Vertrauensbildung

Zusammenfassung: Fokus auf 3 wesentliche Punkte

Prinzip: Je schwerwiegender die Folgen, je mehr menschliche Kontrolle

Prio 1: Vermeidung Eskalation mit existentiellen Risiken	**Keine KI-gesteuerte nukleare Reaktion auf technische Warnung vor RakAngriff**
Prio 2: Pol-Mil. Verantwortung für militär. Operationen	**Keine KI-gest. „autonome“ OpPlanung >Ethisch-polit-militär. Entsch.-Kriterien**
Prio 3: Wahrung Humanit. VR Schutz von Nicht-Kombat.	**Keine KI-gest. „autonome“ Zielauswahl >Unterscheidung – Verhältnismäßigkeit**

Anhänge

- OSZE Wiener Dokument
- OSZE Cyber CBMs

OSZE Wiener Dokument i.d.F. von 2011

Notifikationen + Beobachtungen mil Aktivitäten (WD V, VI)

Notifikation	Pers	KpfPz	gepKpfFz	Artillerie	KpfFlz/KHS
ÜbUmfang	9.000	250	500	250	200 (sorties)
See-/Luftland.	3.000				

Beobachtung	Pers	KpfPz	gepKpfFz	Artillerie	KpfFlz/KHS
ÜbUmfang	13.000	300	500	250	(sorties)
See-/Luftland.	3.500				

Ausnahmen (1) Alarmübungen bis 72 Std Dauer
(2) Parallelübungen, getrennte OpFü, kein gemeinsamer Zweck

Einschrän-kungen (1) Nur Land-/Luft-SK, keine See-SK, LV, stratTr u.a.
(2) *de facto* unwirksam gegenüber nicht-staatl. Akteuren

Krise Freiwillige Beobachtung „ungewöhnlicher mil Aktivitäten“

Einladung zum Besuch v. Militärflughafen und Vorführung neuer Waffensysteme

OSCE Cyber CBMs

gem. PC Dec. 1106/3 Dec **2013** u. PC Dec.1202/10 Mar **2016**

1. Threat information sharing
2. Cross-border co-operation
3. Consultations to reduce misperceptions
4. Open, interoperable, secure, reliable internet
5. OSCE capacity building platform for dialogue and exchange of best practices
6. National legislation to facilitate co-operation
7. Info on nat'l strategies. policies, programs
8. 24/7 points of contact
9. ICT terminology related to security: Info and development common glossary
10. Use of OSCE Communication Network
11. Regular meetings of Informal Working Group at least three times per year

12. Joint action to reduce tensions
13. Effective communication channels to prevent / reduce risks
14. Public-Private-Partnership: best practices on responses to security risks
15. Protect Critical Infrastructure through (sub-)regional co-operation
16. Sharing of vulnerability analysis and information on remedies