الأمم المتحدة

A/80/78*

الجمعية العامة

Distr.: General
5 June 2025
Arabic
Original: Chinese/English/French/
Russian/Spanish

الدورة الثمانون

البند 101 من القائمة الأولية**

نزع السلاح العام الكامل

الذكاء الاصطناعي في المجال العسكري وآثاره على السلام والأمن الدوليين

تقرير الأمين العام

موجز

يقدم هذا التقرير موجزا جامعا لعناصر مستقاة من الردود الواردة من الدول الأعضاء والدول التي لها مركز المراقب عملا بالقرار 239/79 من دون المساس بمواقفها الفردية. وهو يتضمن ما يتيحه الذكاء الاصطناعي في المجال العسكري من فرص ويطرحه من تحديات؛ وقائمة بالمقترحات المعيارية القائمة والناشئة؛ ومسحاً للمبادرات المتعلقة بالذكاء الاصطناعي في المجال العسكري؛ والاعتبارات المتعلقة بالخطوات المقبلة؛ وملاحظات الأمين العام واستنتاجاته.

* أعيد إصدارها لأسباب فنية في 22 كانون الأول/ديسمبر 2025.

** A/80/50

221225 160625 25-06526 (A)

الرجاء إعادة استعمال الورق

المحتويات

الصفحة

أولا – مقدمة 3

ثانيا – معلومات أساسية 3

ثالثا – الفرص والتحديات 4

رابعا – المقترحات المعيارية القائمة والناشئة 7

خامسا – المبادرات المتعلقة بالذكاء الاصطناعي في المجال العسكري 9

سادسا – الخطوات المقبلة 11

سابعا – ملاحظات الأمين العام واستنتاجاته 13

المرفق الأول

الردود الواردة 17

ألف – الدول الأعضاء والدول المراقبة 17

الأرجنتين 17

النمسا 19

شيلي 22

الصين 25

مصر 27

السلفادور 29

فنلندا 32

فرنسا 34

المانيا 37

اليونان 41

الهند 43

إندونيسيا 44

إيران (جمهورية – الإسلامية) 48

إسرائيل 49

إيطاليا 51

اليابان 53
ليتوانيا 57
المكسيك 58
هولندا (مملكة –) 62
نيوزيلندا 66
النرويج 68
باكستان 72
بيرو 77
جمهورية كوريا 80
الاتحاد الروسي 84
صربيا 89
سنغافورة 92
إسبانيا 94
سويسرا 97
أوكرانيا 102
المملكة المتحدة لبريطانيا العظمى وأيرلندا الشمالية 104
باء – الاتحاد الأوروبي 108

المرفق الثاني

Replies received from international and regional organizations, the International Committee of the Red Cross, civil society, the scientific community and industry 111
A. International and regional organizations 111
African Commission on Human and Peoples' Rights 111
B. International Committee of the Red Cross 115
C. Civil society 120
Autonorms 120
Global Commission on Responsible Artificial Intelligence in the Military Domain 123
InterAgency Institute 128
International Committee for Robot Arms Control 130
International Humanitarian Law and Youth Initiative 131
Peace Movement Aotearoa and Stop Killer Robots Aotearoa New Zealand 135

Ploughshares 138
Soka Gakkai International 140
Stop Killer Robots 142
Stop Killer Robots Youth Network 145
Unione degli Scienziati Per Il Disarmo 149
Women's International League for Peace and Freedom 150
D. Scientific community 151
AI, Automated Systems, and Resort-to-Force Decision Making Research Project, the Australian National University 151
Queen Mary University of London, T.M.C. Asser Institute, University of Southern Denmark and University of Utrecht 167
United Nations Institute for Disarmament Research 161
E. Industry 165
Microsoft 165

أولا - مقدمة

1 - طلبت الجمعية العامة من الأمين العام في الفقرة 7 من قرارها 239/79 بشأن الذكاء الاصطناعي في المجال العسكري وآثاره على السلام والأمن الدوليين، أن يلتمس آراء الدول الأعضاء والدول المراقِبة بشأن الفرص والتحديات التي يطرحها تطبيق الذكاء الاصطناعي في المجال العسكري على السلام والأمن الدوليين، مع التركيز بوجه خاص على مجالات غير منظومات الأسلحة الفتاكة الذاتية التشغيل، وأن يقدم تقريرا موضوعيا يلخص تلك الآراء ويصنف المقترحات المعيارية القائمة والناشئة ويتضمن مرفقا يورد هذه الآراء، إلى الجمعية العامة في دورتها الثمانين لمواصلة مناقشته من قبل الدول. وفي الفقرة 8 من القرار نفسه، طلبت الجمعية أيضا إلى الأمين العام أن يلتمس آراء المنظمات الدولية والإقليمية واللجنة الدولية للصليب الأحمر والمجتمع المدني والأوساط العلمية والقطاع المعني تمهيدا لإدراجها باللغة الأصلية التي وردت بها في مرفق التقرير الآنف الذكر. ويُقدَّم هذا التقرير تلبيةً لهذين الطلبين.

2 - وفي 12 شباط/فبراير 2025، أرسل مكتب شؤون نزع السلاح مذكرة شفوية إلى جميع الدول الأعضاء والدول التي لها مركز المراقب وجّه فيها انتباهها إلى الفقرة 7 من قرار الجمعية العامة 239/79، ويلتمس آراءها بشأن المسألة. كما أُرسلت مذكرات شفوية ورسائل إلى الكيانات المحددة في الفقرة 8 من القرار نفسه لاسترعاء انتباهها إلى تلك الفقرة والتماس آرائها بشأن المسألة. وترد في مرفقَي هذا التقرير الآراءُ التي كانت وردت حتى 11 نيسان/أبريل 2025. وستُنشَر أي آراء ترد بعد ذلك التاريخ في الموقع الشبكي لمكتب شؤون نزع السلاح باللغة الأصلية التي قُدِّمت بها.

3 - وتتضمن الفروع الثاني إلى السادس من هذا التقرير موجزا جامعا للعناصر المستقاة من الردود الواردة من الدول الأعضاء والدول التي لها مركز المراقب، من دون المساس بمواقفها الفردية. وترد في الفرع السابع ملاحظات الأمين العام واستنتاجاته.

ثانيا - معلومات أساسية

4 - أشارت الدول إلى أوجه التطور السريعة في العلوم والتكنولوجيا عموما، وفي مجال الذكاء الاصطناعي خصوصا، ملاحظةً تأثيرها الواسع النطاق على المجتمع. وبشكل أكثر تحديدا، أشارت الدول إلى أن لدى الذكاء الاصطناعي القدرة على إدخال تحويل على كل جانب من جوانب الشؤون العسكرية وتأثير كبير على السلام والأمن الدوليين.

5 - وأشار العديد من الدول إلى التطبيقات الحالية للذكاء الاصطناعي في المجال العسكري وكذلك إلى الجهود الخاصة بها لاستخدام الذكاء الاصطناعي في العمليات الدفاعية. ورغم الاعتراف بأهمية المناقشات المتعلقة بمنظومات الأسلحة الفتاكة الذاتية التشغيل، أشارت الدول إلى أن مسألة الذكاء الاصطناعي في المجال العسكري هي أوسع نطاقاً وتشمل مجموعة واسعة من القدرات.

ثالثا - الفرص والتحديات

6 - لوحظ أن الذكاء الاصطناعي يتيح فرصا ويطرح تحديات في آن، لذا ينبغي تناوله بطريقة واقعية. وجرى التسليم بأن سرعة تطور الذكاء الاصطناعي تعني أنه لا يمكن التنبؤ في الوقت الحاضر بمجمل هذه الفرص والتحديات. واقتُرِحَ عدم وصم التكنولوجيا نفسها.

ألف - الفرص المتاحة

7 - أُقِرَّ بعامل السرعة باعتباره ميزة رئيسية للذكاء الاصطناعي، بما في ذلك في تحليل المعلومات واتخاذ القرارات. وأُشيرَ أيضا إلى الحجم باعتباره ميزة هو الآخر، حيث يمكن للذكاء الاصطناعي أن يكون بمثابة "عامل مضاعِف للقوة". وأشارت دول عدة إلى قدرة الذكاء الاصطناعي على تعزيز الكفاءة والصوابية والدقة، ما يؤدي إلى انخفاض احتمال الخطأ مقارنة بالبشر. ومن الخصائص الإضافية التي لوحظت الموثوقية والسلامة والمتانة.

التطبيقات

8 - أشارت دول عدة إلى تطبيقات الذكاء الاصطناعي في مجال الاستخبارات والمراقبة والاستطلاع، حيث يمكن استخدامها لتحليل مجموعات البيانات الكبيرة بكفاءة، وتسهيل كشف التهديدات، وتمكين الإلمام بالحالة السائدة، والقيام بعمليات أكثر دقة. ولوحظ أن هذه الخصائص نفسها مكّنت الذكاء الاصطناعي من دعم عملية صنع القرار والقيادة والتحكم، ما قد يؤدي إلى عمليات أكثر دقة والحد من المخاطر على المدنيين وتوفير قدرٍ أكبر من الحماية للأعيان المدنية. ومع ذلك، جرى التأكيد أيضا على أنه ينبغي لأدوات الذكاء الاصطناعي ألا تحل محل عملية صنع القرار البشرية.

9 - وأشارت دول عدة إلى أنه يمكن دمج الذكاء الاصطناعي في المنظومات المسيَّرة. ولوحظ أنه يمكن للذكاء الاصطناعي أن يحسن التنسيق والتواصل بين الجهات الفاعلة العسكرية وبين الجهات الفاعلة العسكرية وغيرها من الجهات، مثل مقدمي المساعدات الإنسانية. وبصورة عامة، لوحظ أن من شأن الذكاء الاصطناعي أن يقلل من عبء المهام الروتينية أو المتكررة ويزيد القدرات البشرية في المهام المعقدة.

10 - وبحسب بعض الدول، يمكن استخدام الذكاء الاصطناعي لتعزيز أمن تكنولوجيا المعلومات والاتصالات من خلال كشف الاختراقات أو الأنشطة الخبيثة الأخرى، بما في ذلك حماية البنية التحتية الحيوية. ولوحظ أنه يمكن استخدام الذكاء الاصطناعي لكشف المحتوى المُنشأ بواسطة الذكاء الاصطناعي المستخدم في المعلومات المغلوطة والمعلومات المضللة، وكذلك لتحديد خطاب الكراهية أو الدعاية المغرضة أو التغييرات في المشاعر العامة.

11 - وأُشيرَ إلى تطبيقات أخرى للذكاء الاصطناعي لا تتعلق مباشرة بالقتال، بينها الاستخدام الأمثل للوجستيات، والصيانة الوقائية، والمشتريات، وتخصيص الموارد، والإدارة، والمحاكاة، والتدريب.

السلام والأمن الدوليان

12 - اعتبرت دول عدة أن الذكاء الاصطناعي يمكن أن يسهم في صون السلام والأمن الدوليين، فعلى سبيل المثال، يمكن للإلمام بالحالة المدعوم بالذكاء الاصطناعي أن يساعد في التخفيف من المخاطر وأن

يسهم في تخفيف حدة النزاعات. كما لوحظ أن استخدام الذكاء الاصطناعي يمكن أن يقلل من المخاطر التي يتعرض لها الأفراد العسكريون، على سبيل المثال، عن طريق استبدال البشر في تنفيذ مهام خطرة معينة كالتخلص من الذخائر غير المنفجرة، أو عن طريق دعم عمليات البحث والإنقاذ في المواقع النائية.

13 - واقتُرحَ أنه يمكن للذكاء الاصطناعي أن يحسِّن تطبيق القانون الدولي الإنساني، ولا سيما مبادئه الأساسية المتمثلة في التمييز والتناسب واتخاذ التدابير الاحترازية في الهجوم، وحماية المدنيين والأعيان المدنية. وفي هذا الصدد، أشارت دول عدة إلى قدرة الذكاء الاصطناعي على تحسين الإلمام بالحالة عموما، وفهم البيئة المدنية خصوصا، فضلاً عن قدرته على زيادة الدقة والحد من مخاطر الخطأ البشري. وأُشيرَ أيضا إلى أن الذكاء الاصطناعي يمكن أن يسهّل إجراء التحقيقات في إصابات المدنيين ما يضمن تاليا محاسبة المسؤولين عنها.

14 - وأشارت دول عدة إلى أن الذكاء الاصطناعي يمكن أن يساعد في رصد تنفيذ اتفاقات نزع السلاح وعدم الانتشار وتحديد الأسلحة والتحقق منها. وأُشيرَ إلى إمكانية استخدام الذكاء الاصطناعي لدعم بعثات حفظ السلام، بما في ذلك تيسير التخطيط واللوجستيات ورصد وقف إطلاق النار. وشملت التطبيقات الأخرى ذات الصلة بالذكاء الاصطناعي التي جرى تحديدها أمنَ الحدود، ومكافحة الإرهاب، وكشف برامج الأسلحة غير المشروعة، والتقديم الأمثل للمساعدات الإنسانية، والاستجابة للكوارث.

باء - التحديات

15 - أشارت دول عدة إلى أن التطورات السريعة في التكنولوجيات الناشئة عموما والذكاء الاصطناعي خصوصا تطرح تحديات للسلام والأمن الدوليين. ومع أن من المهم فهم هذه التحديات، فإنه يتعذر حاليا التنبؤ بها كلها.

16 - وسُلط الضوء على الهواجس التالية المتعلقة بالذكاء الاصطناعي:

- تسارُع في حلقة المراقبة والتوجيه والتقرير والتصرف، ما يقلّص الوقت المتاح لاتخاذ القرار
- زيادة التشغيل الذاتي وفقدان التحكم البشري، وبخاصة في سياق استخدام القوة
- احتمال إساءة الاستخدام أو الاستخدام الكيدي
- الثقة المفرطة للبشر في تطبيقات الذكاء الاصطناعي
- تعميق أوجه عدم التكافؤ التكنولوجي بين الدول

السلام والأمن الدوليان

17 - أشارت دول عدة إلى أن دمج الذكاء الاصطناعي في المجال العسكري يمكن أن يطرح تحديات للسلام والأمن الدوليين. ويمكن أن يؤدي استخدام الذكاء الاصطناعي إلى زيادة مخاطر سوء الفهم وسوء التقدير والتصعيد غير المقصود، بما في ذلك نتيجةً لتزايُد سرعة وحجم العمليات المدعومة بالذكاء الاصطناعي أو بسبب الأعطال الفنية. ويمكن أن تؤدي هذه الأسباب أيضا إلى خفض عتبة استخدام القوة. وقد أعربت دول عدة عن قلقها إزاء ظهور سباق تسلح في هذا المجال. وأُشير إلى أن استخدام الذكاء

الاصـطناعي يمكن أن يحول التوازن من الأعمال الدفاعية إلى الأعمال الهجومية وإلى أن زيادة الاختلالات بين الدول يمكن أن تؤدي إلى زيادة عدم الاستقرار، وتاليا إلى تقويض السلام والأمن الدوليين.

18 - وأعربت دول عدة عن قلقها إزاء الأثر المحتمل المزعزِع للاســـتقرار لانتشـــار قدرات الذكاء الاصـطناعي، بما في ذلك وصـولها إلى الجهات الفاعلة غير التابعة للدول. ولوحظ أنه لا يوجد حاليا إطار متعدد الأطراف يتيح التحكم بانتشار الأسلحة التي تدمج قدرات الذكاء الاصطناعي.

الاعتبارات التكنولوجية

19 - نظرت الدول في المخاطر الناشئة عن الاعتبارات التكنولوجية، ومن بينها:

- العطل الفني والخلل الفني
- العيوب في التصميم
- السلوك غير المقصود، المنحرف عن معايير التصميم
- قابلية التعرض للهجمات السيبرانية وتسمُّم البيانات
- التحيزات الخوارزمية وتحيزات البيانات، بما في ذلك التحيز الجنساني
- تحيز الأتمتة الناتج عن عدم كفاية تدريب المشغلين البشريين
- مخاوف متعلقة بالخصـوصـية ناشـئة عن جمع ومعاملة كمّ كبير من البيانات الشـخصـية لتدريب نماذج الذكاء الاصطناعي
- المشاكل الناجمة عن سوء تدريب نماذج الذكاء الاصطناعي
- المشاكل الناجمة عن سوء الاختبار والتقييم والتثبت وإجراءات التحقق
- الأخطاء في اختيار الهدف
- الاستهلاك المفرط للطاقة
- الاعتماد المفرط على مقدمي الخدمات الخارجية

20 - وأعربت دول عدة عن قلقها بشأن شفافية قدرات الذكاء الاصطناعي المعقدة وإمكانية شرحها، التي غالبا ما يشـار إليها بـــ "الصـناديق السـوداء". وأُعربَ عن القلق أيضـا إزاء اسـتخدام تطبيقات الذكاء الاصطناعي المدنية، مثل الذكاء الاصطناعي التوليدي الذي يمكن أن يضيف إلى حالة أي نزاع درجة أخرى من التعقيد وعدم اليقين. وأثارت عدة دول مخاوف بشـأن التقارب بين الذكاء الاصـطناعي والتكنولوجيات الأخرى.

الاعتبارات القانونية والإنسانية

21 - لاحظت دول عدة أن الذكاء الاصـطناعي يطرح تحديات أمام التقيد بالقانون الدولي، ولا سـيما القانون الدولي الإنساني والقانون الدولي لحقوق الإنسان. ومن شأن استخدام الذكاء الاصطناعي أن يؤدي إلى استخدام عشوائي للقوة وهو يطرح تساؤلات حول المسؤولية والمساءلة في حالة ارتكاب أفعال غير قانونية

أو خاطئة. ومن القضايا المطروحة ذات الصلة حمايةُ المدنيين والبنية التحتية المدنية، وكذلك احتمال زيادة حدة وفتك النزاعات بالنسبة إلى المقاتلين.

22 - وأثارت عدة دول مخاوف أخلاقية، مشيرةً إلى أن استخدام الذكاء الاصطناعي يمكن أن يقلل إلى أدنى حد من نطاق التعاطف والتفكير الأخلاقي والتقدير البشري.

المجالات المحتملة لإساءة الاستخدام

23 - لاحظت دول عدة إمكانية استخدام الذكاء الاصطناعي في شن هجمات سيبرانية من قبل الدول أو جهات فاعلة غير تابعة للدول، بما في ذلك ضد البنية التحتية الحيوية. ومن المحتمل أيضا أن يُستخدم الذكاء الاصطناعي في حملات المعلومات المغلوطة والمعلومات المضللة، بما في ذلك لإنتاج معلومات كاذبة ومزيفة بشكل عميق، وكذلك للنشر بواسطة الروبوتات القائمة على الذكاء الاصطناعي. إن استخدام هذه المعلومات المغلوطة والمعلومات المضللة، للتأثير على الانتخابات مثلا، يمكن أن يؤدي إلى زعزعة الاستقرار.

أسلحة الدمار الشامل

24 - شددت دول عدة على أهمية الحفاظ على التحكم البشري بالأسلحة النووية ومنظومات إيصالها وأعربت عن قلقها إزاء إمكان دمج الذكاء الاصطناعي في منظومات القيادة والتحكم والاتصالات النووية. وأُشيرَ إلى التزام دول معينة حائزة للأسلحة النووية بالحفاظ على التحكم والمشاركة البشريين في كل الإجراءات البالغة الأهمية في إبلاغ وتنفيذ القرارات السيادية المتعلقة باستخدام الأسلحة النووية. وأُشيرَ إلى التداعيات المحتملة على الاستقرار الاستراتيجي والتصعيد.

25 - وأعربت دول عدة عن قلقها من أن الذكاء الاصطناعي يمكن أن يسهِّل انتشارَ أسلحة الدمار الشامل، بما في ذلك إلى جهات فاعلة غير تابعة للدول. وفي هذا السياق، كانت إحدى الهواجس الخاصة هي إمكان استخدام الذكاء الاصطناعي لتطوير وإنتاج أسلحة بيولوجية. وجرى التشديد على أنه بموجب أحكام المعاهدات القائمة، يجب عدم استخدام الذكاء الاصطناعي لهذه الغاية. كما أُعرب عن رأي مفاده أنه يمكن استخدام الذكاء الاصطناعي للجم انتشار أسلحة الدمار الشامل.

رابعا - المقترحات المعيارية القائمة والناشئة

26 - أعربت دول عدة عن ضرورة استخدام الذكاء الاصطناعي للأغراض السلمية، بما في ذلك التسوية السلمية للمنازعات. كما شددت الدول على أهمية التصدي للمخاطر الناشئة عن الذكاء الاصطناعي في المجال العسكري والحد منها، حيث أشار البعض إلى أنه ينبغي التصدي بشكل جماعي للتحديات الناشئة عن الذكاء الاصطناعي العسكري.

27 - وعند تناوُل الذكاء الاصطناعي في المجال العسكري، دعت الدولُ إلى اتباع نهجٍ يتسم بما يلي:

- المرونة والتوازن والواقعية والتدرج، وبالتالي القدرة على التكيف مع أوجه التقدم التكنولوجي
- الطابع الاحترازي

- التركيز على كامل دورة عمر الذكاء الاصطناعي، بما في ذلك مرحلة ما قبل التصميم والتصميم والتطوير والتقييم والاختبار والنشر والاستخدام والبيع والشراء والتشغيل وإيقاف التشغيل
- الاستناد إلى تطبيقات واستخدام الذكاء الاصطناعي، عوض الاستناد إلى التكنولوجيا نفسها
- أن يكون انعكاسا للالتزامات القائمة

واقتُرِحَ أن تميز الجهودُ المبذولة في هذا المجال تمييزا واضحا بين أوجه الاستخدام الفتاكة وغير الفتاكة.

الاعتبارات القانونية

28 - أشارت الدول إلى القرار 239/79 الذي أكدت فيه الجمعية العامة أن القانون الدولي، بما في ذلك ميثاق الأمم المتحدة والقانون الدولي الإنساني والقانون الدولي لحقوق الإنسان، ينطبق على ما يخضع لأحكامه من مسائل تطرأ في جميع مراحل دورة عمر الذكاء الاصطناعي، بما في ذلك المنظومات التي تستمد فعاليتها من الذكاء الاصطناعي، في المجال العسكري. ولوحظ أن القانون الدولي عموما، والقانون الدولي الإنساني خصوصا، لا يحظّران بشكل قاطع استخدام قدرات الذكاء الاصطناعي.

29 - وأكدت الدول أنها ممتثلة للقانون الدولي في استخدامها للذكاء الاصطناعي في المجال العسكري. واقتُرح أن يكون الامتثال للالتزامات القانونية، لا سيما تلك المستمدة من القانون الدولي، أحد الاعتبارات الرئيسية الواجبة في إدارة الذكاء الاصطناعي وتصميمه ونشره في المجال العسكري. وعلاوة على ذلك، أُعربَ عن رأي مفاده أنه ينبغي تصميم الذكاء الاصطناعي على نحو يعزز الامتثال للقانون الدولي الإنساني. وشددت دول عدة على أهمية إجراء استعراضات قانونية للأسلحة أو وسائل أو أساليب الحرب الجديدة في هذا الصدد.

30 - وشددت دول عدة على أهمية مراعاة الاعتبارات الأخلاقية، إضافة إلى الأطر القانونية.

الاعتبارات المتعلقة بالسلام والأمن الدوليين

31 - أشارت دول عدة إلى أنه ينبغي للذكاء الاصطناعي في المجال العسكري أن يعزز السلام والأمن الدوليين وأن يُستخدم بطريقة لا تؤدي إلى عدم الاستقرار أو التصعيد. وأُعرب عن رأي مفاده أنه ينبغي للدول أن تمتنع عن السعي إلى تحقيق ميزة عسكرية مطلقة من خلال الذكاء الاصطناعي، وينبغي لها أن تضمن ألا تصبح هذه التكنولوجيا أداة لشن غزو أو للسعي إلى الهيمنة.

32 - ورأت دول عدة أنه ينبغي للذكاء الاصطناعي ألا يقوض ما هو قائم من اتفاقات لنزع السلاح وعدم الانتشار وتحديد الأسلحة. وأُطلقت دعوات لبذل جهود في سبيل منع انتشار تكنولوجيا الذكاء الاصطناعي إلى جهات غير تابعة للدول. وجرى التشديد على أهمية تجنب آليات الرقابة الدولية التعسفية أو الرقابة التمييزية على الصادرات.

الاستخدام المسؤول للذكاء الاصطناعي في المجال العسكري

33 - رأت دول عدة أنه ينبغي تطبيق الذكاء الاصطناعي بطريقة مسؤولة طوال دورة عمره. وأُعربَ عن رأي مفاده أنه ينبغي ربط مفهوم المسؤولية بالمشروعية والمساءلة.

34 - وشددت دول عدة على أهمية اتباع نهج يركز على الإنسان إزاء الذكاء الاصطناعي. وشدد العديد من الدول على أهمية ممارسة التحكم والمسؤولية البشريين في جميع الأوقات. وأُشيرَ إلى أهمية مفاهيم من قبيل "التحكم والحُكم البشريان المناسبان للسياق" و "التحكم البشري المجدي". وعلى النقيض من ذلك، وبحسب دول أخرى، لم تعرَّف هذه المفاهيم بشكل كافٍ. وأُعربَ عن رأي مفاده أن استخدام مفهوم "التحكم البشري المجدي" يمكن أن يعوق البحث المشروع أو يقيد من دون مبرر استخدامَ الذكاء الاصطناعي في المجال العسكري.

35 - وشددت الدول على أهمية ضمان المسؤولية والمساءلة البشريتين، بما في ذلك في إطار سلسلة مسؤولة تتولى القيادة والتحكم البشريين، وفقاً للقانون الدولي.

الاعتبارات التكنولوجية

36 - نظرت الدول في مبادئ الحوكمة من منظور تكنولوجي، بما في ذلك ما يلي:

- الأمن، لضمان صلابة منظومات الذكاء الاصطناعي في وجه التهديدات الخارجية
- السلامة، بما في ذلك من خلال دمج ضوابط للحد من الضرر إلى أدنى حد
- الموثوقية، لمنع حصول عواقب وأعطال غير مقصودة
- حدود وقيود عملياتية واضحة لمنع السلوك غير المقصود
- حالات استخدام واضحة تماما
- إمكانية التحكّم، عبر ضمان إتاحة التفاعل المناسب بين الإنسان والآلة والتقليل من التحيز
- العدالة والإنصاف
- حماية الخصوصية
- قابلية الشرح والفهم والتتبع
- الشفافية

37 - وسُلط الضوء على استخدام بيانات التدريب التي تتيح الامتثال التام للقانون الدولي. وأكدت دول عدة على أهمية الاختبار طوال دورة العمر للكشف عن الأخطاء وضمان الموثوقية. وجرى التشديد على ضرورة توفير التدريب الكافي للموظفين العاملين الذين يستخدمون الذكاء الاصطناعي للتخفيف من المخاطر وضمان الامتثال للقانون الدولي الإنساني. وتم التأكيد على أهمية مراقبة أداء النظام طوال دورة عمره، وعلى أهمية اتخاذ تدابير لتعطيل الأنظمة بشكل آمن عند توقفها عن العمل.

خامسا - المبادرات المتعلقة بالذكاء الاصطناعي في المجال العسكري

المنتديات الدولية

38 - أشارت دول عدة إلى المناقشات الجارية في الأمم المتحدة، وكذلك إلى ميثاق المستقبل (قرار الجمعية العامة 1/79) والتعاهد الرقمي العالمي المرفق به، وقرار الجمعية بشأن الذكاء الاصطناعي في

المجال العسكري وآثاره على السلام والأمن الدوليين (القرار 239/79). وأشيرَ أيضا إلى اجتماع مجلس الأمن المعقود بصيغة آريا بشأن تسخير الذكاء الاصطناعي الآمن والشامل للجميع والموثوق من أجل صون السلام والأمن الدوليين، المعقود في 4 نيسان/أبريل 2025.

39 - وأشارت الدول إلى المناقشات المتعددة الأطراف بشأن المواضيع المتعلقة بالذكاء الاصطناعي في المجال العسكري، مثل المناقشات التي أُجريت في هيئة نزع السلاح في إطار بند جدول الأعمال المعنون "توصيات بشأن التفاهمات المشتركة المتصلة بالتكنولوجيات الناشئة في سياق الأمن الدولي"، وأعمال فريق الخبراء الحكوميين المعني بالتكنولوجيات الناشئة في مجال منظومات الأسلحة الفتاكة الذاتية التشغيل، المُنشأ بموجب اتفاقية حظر أو تقييد استعمال أسلحة تقليدية معينة يمكن اعتبارها مفرطة الضرر أو عشوائية الأثر، وكذلك قرارات الجمعية العامة بشأن منظومات الأسلحة الفتاكة الذاتية التشغيل (القراران 241/78 و 62/79).

40 - وأشارت الدول أيضا إلى مشاركتها في الأنشطة المتصلة بالذكاء الاصطناعي في المجال العسكري التي ينظمها مكتب شؤون نزع السلاح ومعهد الأمم المتحدة لبحوث نزع السلاح.

المبادرات التي تقودها الدولة

41 - أشارت دول عدة إلى أنها اتخذت مبادرات تتعلق بالذكاء الاصطناعي في المجال العسكري أو شاركت فيها، بينها:

- الإعلان السياسي بشأن الاستخدام العسكري المسؤول للذكاء الاصطناعي وخاصية التشغيل الذاتي وعملية تنفيذه اللاحقة
- عملية الذكاء الاصطناعي المسؤول في المجال العسكري، والتي شملت عقد مؤتمرات في مملكة هولندا عام 2023، حيث تم تأييد "دعوة إلى العمل"، وفي جمهورية كوريا عام 2024، حيث تم تأييد "مخطط للعمل". وكان من المتوقع أن تنشر اللجنة العالمية المعنية بالذكاء الاصطناعي المسؤول في المجال العسكري تقريراً قبل انعقاد المؤتمر المقبل المزمع عقده في إسبانيا عام 2025
- قمة عمل الذكاء الاصطناعي المعقودة في فرنسا عام 2025، والتي اعتُمد فيها إعلان باريس بشأن الحفاظ على التحكم البشري في منظومات الأسلحة القائمة على الذكاء الاصطناعي
- مؤتمر قمة السلامة في مجال الذكاء الاصطناعي المعقود في المملكة المتحدة لبريطانيا العظمى وأيرلندا الشمالية عام 2023، الذي اعتُمد فيه إعلان بلتشلي
- العمل في مجال الذكاء الاصطناعي الجاري في سياق مجموعة الدول السبع
- الشراكة الدفاعية في مجال الذكاء الاصطناعي
- مبادرة الحوكمة العالمية للذكاء الاصطناعي المقترحة في عام 2023

42 - وأُعرب عن رأي مفاده أن هذه المبادرات، على فائدتها، يمكن أن تؤدي إلى التجزئة. كما أُعرب عن القلق من أن نتائج هذه المبادرات لم تراعِ آراء جميع الدول المعنية ويمكن أن تقوض العمل الشامل للجميع في هذا المجال.

المبادرات الإقليمية

43 - أشارت الدول إلى أهمية المبادرات الإقليمية لتعزيز المناقشات الجامعة والمحدَّدة السياق بشأن الذكاء الاصطناعي في المجال العسكري. وتشمل الأمثلة في هذا الصدد ما يلي:

- البيان المشترك بشأن التعاون في مجال الذكاء الاصطناعي في قطاع الدفاع، الذي اعتُمدَ في خلوة اجتماع وزراء دفاع رابطة أمم جنوب شرق آسيا في عام 2025
- المؤتمر السادس عشر لوزراء الدفاع في الأمريكتين، المعقود في عام 2024، الذي اعتُمد فيه إعلان مندوسا
- الأنشطة المضطلَع بها في سياق منظمة حلف شمال الأطلسي، بما في ذلك استراتيجيته المتعلقة بالذكاء الاصطناعي، التي نُقحت آخر مرة في عام 2024، ومبادئ الاستخدام المسؤول، التي وُضعت في عام 2021
- المشاورات الإقليمية في سياق عملية الذكاء الاصطناعي المسؤول في المجال العسكري في عام 2024، التي أُجريت في تركيا وسنغافورة وشيلي وكينيا و(مملكة -) هولندا عام 2024

المبادرات المحلية

44 - أشارت الدول إلى جهودها المحلية، بما في ذلك التشريعات والقواعد التنظيمية والاستراتيجيات والهيئات القائمة في مجال الذكاء الاصطناعي، فضلاً عن الجهود المبذولة لتطويرها.

سادسا - الخطوات المقبلة

45 - دعت الدولُ إلى إجراء حوار حول الذكاء الاصطناعي في المجال العسكري. ودعت دول عدة إلى إجراء مزيد من الدراسة لأثر الذكاء الاصطناعي في المجال العسكري على السلام والأمن الدوليين.

46 - ولاحظت دول عدة أن الهدف من مواصلة الحوار ينبغي أن يكون التخفيف من المخاطر التي يشكلها الذكاء الاصطناعي في المجال العسكري. واقتُرح أن يكون الهدفُ من الحوار وضعَ أطر ناظمة أو أطر للحوكمة. ودعت دول عدة إلى وضع معايير وقواعد ومبادئ تحكم دورة عمر الذكاء الاصطناعي في المجال العسكري. وفي حين أعربت بعض الدول عن تفضيلها لوضع إطار ملزم قانوناً، اعتبرت دول أخرى أن اعتماد تدابير قانونية جديدة غير ضروري في الوقت الحاضر. كما أُعربَ عن رأي مفاده أن المعايير والقواعد والمبادئ يمكن أن تشكل الأساس للالتزامات القانونية في مرحلة لاحقة. وأعربت دول عدة عن معارضتها لمفهوم المعايير والقواعد والمبادئ المتعلقة بالتطوير أو النشر أو الاستخدام المسؤول، مشيرة إلى أن هذا المفهوم لا يحظى بالتوافق. وأُعرب عن رأي مفاده أنه ينبغي تجنب اعتماد قواعد تنظيمية في وقت سابق لأوانه.

47 - وجرى التأكيد على أهمية تجنب الازدواجية والتجزئة في الحوكمة. واعتُبر أنه ينبغي لأي مناقشة بشأن الحوكمة أن تراعي الاعتبارات الإنسانية والأمنية والإنمائية بطريقة متوازنة. وسلطت الدول الضوء على أهمية تجنب القيود التي من شأنها أن تعرقل الابتكار المشروع والتقدم التكنولوجي. ورأت دول عدة أنه لا ينبغي إعاقة استخدامات الذكاء الاصطناعي للأغراض السلمية، ولا سيما من قبل الدول النامية.

48 - وأشـــيرَ إلى أن اتّباع أي نهج إزاء الحوكمة ينبغي أن يأخذ في الاعتبار المراحلَ المختلفة التي بلغتها الدول في دمج الذكاء الاصـطناعي في القدرات العسـكرية وتفاوُتَ بيئاتها الأمنية. وجرى التأكيد على أهمية مشـاركة جميع الدول في المناقشـات بشـأن حوكمة الذكاء الاصـطناعي في المجال العسـكري. ورأى العديد من الدول أنه ينبغي للمناقشـــات المقبلة أن تتبع نهجا قائما على جهات معنية متعددة، بما فيها المنظمات الدولية والإقليمية، والمجتمع المدني، والأوسـاط العلمية، والقطاع الصـناعي. بيد أنه جرى التأكيد على أن اتخاذ القرار ينبغي أن يظل من اختصاص الدول وحدها.

49 - ونظرت الدول في أولويات مختلفة للحوار المقبل حول الذكاء الاصـطناعي في المجال العسـكري، بما في ذلك:

- ضمان الامتثال للقانون الدولي، ولا سيما القانون الدولي الإنساني
- حماية الكرامة الإنسانية وحقوق الإنسان
- البحث عن تفاهمات مشتركة حول التعريفات والمصطلحات
- النظر في تدابير كفالة الشفافية وبناء الثقة
- تناوُل مسألة التشغيل الذاتي في استخدام القوة
- تناوُل مسألة منظومات الذكاء الاصطناعي التي تدعم العمليات القتالية بشكل مباشر
- ضمان وجود آليات ملائمة لحوكمة البيانات
- تعزيز التعاون والمساعدة الدوليين
- دعم بناء القدرات، بما في ذلك من خلال تبادل المعارف ونقل التكنولوجيا وتبادل الممارســـات الجيدة، بغية سد الفجوة الرقمية وفجوة الذكاء الاصطناعي
- التشجيع على مواصلة الحوار الإقليمي
- التشـجيع على وضـع قواعد تنظيمية وطنية، بما في ذلك من أجل ضـمان امتثال القطاع الخاص للقانون الدولي

50 - ورأت دول عدة أن النظر في منظومات الأسلحة الفتاكة الذاتية التشغيل ينبغي أن يشكل جزءاً من أي مناقشة للذكاء الاصطناعي في المجال العسكري. وأُعربَ عن رأي مفاده أن المناقشات الجارية بشأن هذه المنظومات مكمِّلة للمناقشات بشأن الذكاء الاصطناعي في المجال العسكري. وذكَرت دولٌ عدة بمواقفها من منظومات الأسـلحة الفتاكة الذاتية التشـغيل[1]. وفي حين رأى بعضـها أن فريق الخبراء الحكوميين المنشـأ بموجب الاتفاقية المتعلقة بأسلحة تقليدية معينة هو المحفل الأمثل للمناقشات بشأن الذكاء الاصـطناعي في المجال العسكري، رأى بعضُها الآخر أن هذا الفريق ليس المحفل المناسب لمثل هذه المناقشات نظراً لولايته المحددة وعضويته غير العالمية.

51 - ودعا العديد من الدول إلى إجراء مناقشات بشأن الذكاء الاصـطناعي في المجال العسـكري داخل منتديات الأمم المتحدة. واقتُرح أن يشـكل هذا التقرير الأسـاس لهذه المناقشـات. وأشـارت دول إلى أن

(1) لمزيد من التفاصيل، انظر A/79/88.

المناقشات المقبلة ينبغي أن تكون مكمِّلة للعمليات الجارية، مثل الفريق العامل المفتوح العضوية المعني بأمن تكنولوجيا المعلومات والاتصالات واستخدامها.

52 - ورأت دول عدة أن آلية الأمم المتحدة لنزع السلاح تمثل منبراً فعالاً وجامعا وينبغي أن تؤدي دوراً محورياً في المناقشات المقبلة بشأن الذكاء الاصطناعي في المجال العسكري. واقتُرح أن يناقش مؤتمرُ نزع السلاح مسألةَ الذكاء الاصطناعي، ولا سيما من حيث صِلته بالأسلحة النووية. وأُعرب عن رأي مفاده أنه يمكن أيضا إجراء مناقشات في اللجنة الأولى للجمعية العامة، التي يمكن أن تكلّف الأمين العام بتقديم تقارير منتظمة عن حالة التطور التكنولوجي للذكاء الاصطناعي في المجال العسكري. واقترحت دول عدة إجراء مناقشات في سياق هيئة نزع السلاح.

53 - واقتُرح أيضاً أن تُجرى المناقشات في سياق مجلس الأمن.

54 - واقترحت دول عدة إنشاء آلية مخصصة، كفريقٍ عامل مفتوح العضوية، لهذا الأمر. وأُعرب أيضا عن رأي مفاده أن استحداث آلية داخل الأمم المتحدة لن يكون مناسبا في هذا الوقت. ورأى البعض أن أي آلية للأمم المتحدة بشأن هذا الموضوع ينبغي أن تسترشد بتوافق الآراء.

سابعا - ملاحظات الأمين العام واستنتاجاته

55 - للذكاء الاصطناعي القدرة على التأثير على كل جانب من جوانب حياتنا. وعند استخدامه للأغراض السلمية، يمكن أن يؤدي دورا هاما في تيسير تحقيق الالتزامات والأهداف الإنمائية، بما فيها أهداف التنمية المستدامة.

56 - وفي المجال العسكري، يمكن للذكاء الاصطناعي أن يحقق فوائد للقوات العسكرية التي تستخدمه وللسكان المدنيين على حد سواء، من خلال زيادة دقة العمليات وتقليص هامش الخطأ البشري. وفي الوقت نفسه، يطرح الذكاء الاصطناعي في المجال العسكري تحديات خطيرة، أهمها الإبقاء على المسؤولية والمساءلة البشريتين.

57 - إن تأكيد الجمعية العامة في قرارها 239/79 على أن القانونَ الدولي، بما في ذلك ميثاق الأمم المتحدة والقانون الدولي الإنساني والقانون الدولي لحقوق الإنسان، واجبُ التطبيق طوال دورة عمر الذكاء الاصطناعي، يشكل منطلقا أساسيا هاما. ومع ذلك، لا تزال هناك أسئلة هامة حول كيفية تطبيق القانون لم تُحَل بعد.

58 - ويتطلب استخدام الذكاء الاصطناعي العسكري في الحالات التي تنطوي على استخدام القوة انتباها خاصا. ورغم الفوائد التي يمكن تحقيقها لحماية المدنيين والمقاتلين، فإن الاستخدامات المبلَّغ عنها للذكاء الاصطناعي في النزاعات الحالية تثير هواجس بشأن التحكم البشري ودور الذكاء الاصطناعي في تسهيل الأعمال العدائية في المناطق المكتظة بالسكان. فالآلات التي تمتلك القدرة والسلطة التقديرية لإزهاق الأرواح البشرية هي آلات غير مقبولة سياسيا وبغيضة أخلاقيا.

59 - ولن يتم القضاء على المخاطر التي تشكلها الأسلحة النووية ما لم تتم إزالة تلك الأسلحة نفسها. وإلى أن تتحقق الإزالة التامة للأسلحة النووية، أحث جميع الدول التي تمتلك هذه الأسلحة على الاتفاق على أن يكون أي قرار بشأن استخدام الأسلحة النووية بيد البشر وليس عائدا للآلات.

60 - ويمكن للذكاء الاصطناعي أن يخفض الحاجز أمام الجهات الفاعلة غير التابعة للدول في تطوير أو حيازة الأسلحة الكيميائية والبيولوجية. لذا، أحث الدول على الوفاء الكامل بالتزاماتها بموجب الأطر ذات الصلة لنزع السلاح وعدم الانتشار وتحديد الأسلحة وأن تجري تقييما منهجيا وتكون مستعدة بشكل جيد للتصدي للتحديات ولتأثير الذكاء الاصطناعي على تلك الأطر.

61 - ويشكل الدمج المحتمل لتطبيقات الذكاء الاصطناعي المدنية في المجال العسكري مبعثا متزايداً على القلق. إذ إن طابع تكنولوجيات الذكاء الاصطناعي القابل بحكم تكوينه لإعادة الاستخدام لغرض آخر يطرح تحديات للرقابة والشفافية والمساءلة. وأحث الدول على أن تدرس بتمعّن الخطوط المتداخلة بين أوجه التطور في تطبيقات الذكاء الاصطناعي المدنية واستخدامها المحتمل في المجال العسكري.

62 - إن متابعة آليات تعاون إضافية في مجال الذكاء الاصطناعي، وخصوصا على الصعيدين الإقليمي ودون الإقليمي تنطوي على قيمة كبيرة. والمنظمات الإقليمية ودون الإقليمية مجهزة بشكل فريد لوضع وتنفيذ تدابير الشفافية وبناء الثقة كوسيلة للتخفيف من المخاطر. لذا فإنني أشجع الدول على أن تنظر في وضع تدابير الشفافية وبناء الثقة على الصعيدين الإقليمي ودون الإقليمي بما يتناسب مع الخصائص والتحديات الفريدة للذكاء الاصطناعي.

63 - إن المناقشات الجامعة بشأن استخدامات الذكاء الاصطناعي للأغراض السلمية وحوكمته لصالح البشرية متواصلة تحت رعاية الأمم المتحدة، لا سيما في سياق تنفيذ التعاهد الرقمي العالمي. ومع ذلك، فإن نظر الدول الأعضاء في مسألة الذكاء الاصطناعي في المجال العسكري قد تم إلى حد كبير خارج محافل الأمم المتحدة. ويشكل قرار الجمعية العامة 239/79 والتقرير الحالي خطوتين أوليين ملحوظتين في طرح هذه المناقشة الهامة على الأمم المتحدة. وإنني أشجع الدول على إجراء هذه المداولات بطريقة جامعة وبناءة بهدف النهوض بالتفاهمات وتعزيز التعاون الدولي للتخفيف من المخاطر.

64 - والدول مدعوة بقوة إلى استكشاف الجهود، بما في ذلك بناء القدرات، لضمان مشاركة مجدية لجميع الدول في عمليات الأمم المتحدة بشأن هذا الموضوع، وهو أمر ضروري لتعزيز الفهم المشترك، ووضع نهُج مشتركة، والتخفيف من المخاطر المحتملة.

65 - لقد أثبتت الجمعية العامة أنها بارعة في تفويض العمليات التي تعزز المناقشات الشاملة للجميع بشأن القضايا المتعلقة بالتكنولوجيات الناشئة والأمن الدولي، مع تشجيع الإسهامات من الجهات المعنية، بما فيها المنظمات الدولية والإقليمية والمجتمع المدني والأوساط العلمية والقطاع الصناعي. ويتسم هذا النهج المتعدد الجهات المعنية بأهمية خاصة في مجال الذكاء الاصطناعي، فالابتكار مدفوع إلى حد كبير من قبل القطاع الخاص، كما أن الكثير من الخبرات موجود خارج الحكومات، في الأوساط الأكاديمية والعلمية.

66 - **وعليه، أوصي بأن تدرس الدول الأفكار الواردة في هذا التقرير، وأن تقوم، في الدورة الثمانين للجمعية العامة، بخطوات محددة بهدف إطلاق آلية مخصصة وجامعة تتناول بطريقة شاملة مسألة الذكاء الاصطناعي في المجال العسكري وآثاره على السلام والأمن الدوليين.**

المرفق الأول

الردود الواردة

ألف - الدول الأعضاء والدول المراقبة

الأرجنتين

[الأصل: بالإسبانية]
[10 نيسان/أبريل 2025]

يُقدَّم التقرير التالي في ما يتعلق بالقرار 239/79، المعنون "الذكاء الاصطناعي في المجال العسكري وآثاره على السلام والأمن الدوليين"، الذي اتخذته الجمعية العامة في 24 كانون الأول/ديسمبر 2024.

النهج العام

تدرك جمهورية الأرجنتين أن لنشر الذكاء الاصطناعي في البيئة العسكرية آثارا استراتيجية كبيرة. فاستخدام الذكاء الاصطناعي يؤدي إلى تحقيق فوائد ملموسة لوظائف مختلفة غير فتاكة، بيد أنه يحمل معه في الوقت نفسه مخاطر تستدعي استجابةً من منظور القانون الدولي والأخلاقيات والمسؤولية العملياتية. وفي هذا السياق، يجب التمسك باحترام القانون الدولي الإنساني وحقوق الإنسان عند تطوير هذه التكنولوجيات واستخدامها، ويجب أن يكون هناك ضمانات بالحفاظ في جميع الأوقات على المسؤولية البشرية عن اتخاذ القرارات الحاسمة والتحكم بها.

الفرص

يعد الذكاء الاصطناعي، ولا سيما تطبيقاته غير الفتاكة، أداة مشروعة وقيّمة لتعزيز القدرات الدفاعية الوطنية. وتشمل الاستخدامات ذات الأولوية ما يلي:

- تحقيق الاستخدام الأمثل من الناحيتين اللوجستية والتشغيلية
- دعم معالجة المعلومات الاستخبارية
- تعزيز الدفاع السيبراني
- المحاكاة والتدريب والتخطيط الاستراتيجي

وتسهم هذه القدرات في تنفيذ عمليات أكثر كفاءة وأمانًا تحقق تكيفا أفضل مع السيناريوهات الحالية، مما يعزز الفعالية الدفاعية دون تقويض المبادئ الإنسانية أو الالتزامات الدولية للدولة.

التحديات

يطرح التطوير السريع للذكاء الاصطناعي في البيئات العسكرية تحديات يجب التصدي لها بشكل جماعي، بما في ذلك ما يلي:

- خفض عتبة استخدام القوة، وتقصير الإطار الزمني لاتخاذ القرار البشري
- إمكانية وجود تحيز خوارزمي غير مكتشف
- انتشار منظومات ذاتية التشغيل لتصل إلى جهات فاعلة من غير الدول
- خطر ترسيخ أوجه عدم التناظر التكنولوجي بين الدول

وتؤكد هذه المخاطر على الحاجة إلى وضع مبادئ مشتركة وضمانات يمكن التحقق منها وأطر للتعاون.

الحوكمة والتعاون الدولي والإدماج التكنولوجي

من المفهوم لدينا أن أي عملية لوضع السياسات في هذا المجال يجب أن تستند إلى المبادئ التالية:

- ينبغي تجنب اللوائح العامة أو السابقة لأوانها التي تحد من التطوير المستقل للتقنيات الدفاعية المشروعة.
- يجب التمييز بوضوح بين الاستخدامات الفتاكة وغير الفتاكة.
- يجب ضمان رقابة بشرية كبيرة كشرط تشغيلي وسياساتي أساسي.
- ينبغي تعزيز التعاون الدولي الشامل الذي يركز على تعزيز القدرات وسد الفجوات التكنولوجية بين الدول.

وقد أعادت الأرجنتين التأكيد على هذه المبادئ في المنتديات المتعددة الأطراف التي عُقدت مؤخرًا، مؤكدةً على أهمية العمل على وضع معايير مشتركة للاستخدام المسؤول للذكاء الاصطناعي في المجال العسكري، لا سيما فيما يتعلق بالدفاع السيبراني والأمن السيبراني.

ولإعطاء مثال على مبادرة على المستوى الإقليمي، اجتمع الفريق العامل المعني بتطوير الذكاء الاصطناعي وتطبيقه وإدارته على نحو مسؤول في المجال العسكري، في إطار المؤتمر السادس عشر لوزراء الدفاع في الأمريكتين الذي عقد في مندوزا بالأرجنتين في عام 2024، للعمل بشكل جماعي على وضع معايير دولية.

الإشارة إلى ميثاق المستقبل

اسمحوا لي أخيرا أن أسجل في المحضر أن جمهورية الأرجنتين قد نأت بنفسها رسميا عن ميثاق المستقبل، الذي ورد ذكره في ديباجة قرار الجمعية العامة 239/79. وبالتالي، فإن الإشارة إلى الميثاق لا تعني التزام الدولة الأرجنتينية بالميثاق أو تقيدها به أو دعمها له.

النمسا

[الأصل: بالإنكليزية]
[11 نيسان/أبريل 2025]

عملاً بالطلب الوارد في الفقرة 7 من قرار الجمعية العامة 239/79، تود النمسا أن تقدم الأفكار والملاحظات التالية على أساس وطني.

الذكاء الاصطناعي المتصل بالأمن السيبراني والدفاع السيبراني

إن برامجيات الأمن السيبراني القائمة على الذكاء الاصطناعي تُستخدم أصلاً على نطاق واسع للمساعدة في كشف الاختراقات وغيرها من الأنشطة الضارة في شبكات الحواسيب. ومن المرجح أن تتيح أدوات الذكاء الاصطناعي هذه توفير الحماية المؤتمتة بشكل متزايد لنظم تكنولوجيا المعلومات من خلال البحث عن مواطن الضعف والأنشطة المشبوهة بهدف زيادة قدرة البرامجيات والأجهزة على مواجهتها.

وفي الوقت نفسه، يتزايد استخدام أدوات الذكاء الاصطناعي لتحسين مستوى تطور الهجمات السيبرانية وابتكار فيروسات حاسوبية جديدة في سباق بين نماذج الذكاء الاصطناعي الهجومية والدفاعية في مجال الأمن السيبراني. وإضافة إلى ذلك، تعمل البرامجيات القائمة على الذكاء الاصطناعي، بما في ذلك نماذج اللغات الحاسوبية الكبيرة، على خفض حاجز الدخول للجهات الفاعلة الخبيثة التي يمكنها بشكل متزايد إنشاء برامجيات خبيثة دونما حاجة إلى مهارات برمجة واسعة.

الذكاء الاصطناعي المتصل بحملات التضليل الإعلامي كعنصر من عناصر الاستراتيجيات المختلطة

يتزايد استخدام البرامجيات القائمة على الذكاء الاصطناعي القادرة على إنشاء محتوى مزيف ونشره دعماً لحملات التضليل الإعلامي. ومن الأساليب المستخدمة استعمال الذكاء الاصطناعي التوليدي لإنشاء محتوى مخصص ومحلّي على نطاق واسع. وعلاوة على ذلك، فإن برامج الصوت والفيديو المزيفة التي تعتمد على الذكاء الاصطناعي آخذة بالتحسن بشكل سريع، وهي تُستخدم بالفعل على نطاق واسع. ويمكن نشر هذا المحتوى المزيف باستخدام شبكات ضخمة من روبوتات وسائل التواصل الاجتماعي التي تعتمد على الذكاء الاصطناعي لخلق مظهر حصول تغيير في الرأي العام. وبالتالي، فإن الذكاء الاصطناعي يقلل من العوائق التي تحول دون شن حملات تضليل إعلامي واسعة النطاق لأن حجم ونوعية المحتوى المزيف الذي يتم إنشاؤه لم يعد محدودا بعدد أو مهارات المشغلين البشريين.

ومع ذلك، يمكن أيضا استخدام خوارزميات الذكاء الاصطناعي لكشف المحتوى الذي يولده الذكاء الاصطناعي وحملات الدعاية المغرضة، في حين يمكن كشف الصوت والفيديو المزيفين الحقيقيين المخادعين باستخدام أدوات الذكاء الاصطناعي المتخصصة. ومن الضروري اللجوء إلى هذه الأدوات التي تعتمد على الذكاء الاصطناعي لمواجهة الآثار السيئة للذكاء الاصطناعي المستخدم لغرض حملات التضليل.

الذكاء الاصطناعي المتصل بانتشار الأسلحة

يمكن للذكاء الاصطناعي أن يقلص من العائق الذي يحول دون الحصول على الأسلحة، بما فيها أسلحة الدمار الشامل. ونظراً لقدرتها على توفير الخبرة بضغطة زر، فإن نماذج اللغات الحاسوبية الكبيرة والتطبيقات المستندة إليها يمكن أن تسهّل على الجهات الخبيثة تصنيع الأسلحة. وتتراوح حالات الاستخدام من الوصول إلى مخططات التصميم أو طباعة المكونات للأسلحة الصغيرة والأسلحة الخفيفة، إلى تعديل العوامل المُمْرِضة لأغراض الحرب البيولوجية. وإذا ما كانت المعارف المتاحة بسهولة تقلل من نطاق وحجم برامج الأسلحة، فسيكون من الصعب اكتشاف هذه التهديدات ومنعها والتأهب لمواجهتها.

وفي الوقت نفسه، يمكن أيضا استخدام خوارزميات التعلم الآلي لمكافحة انتشار الأسلحة. ونظراً لقدراتها على اكتشاف الحالات الشاذة وتقنيات التعرف على الأشكال، فيمكنها أن تساعد في تحديد الأنشطة الخبيثة، بما في ذلك من خلال كشف التدفقات المالية غير المشروعة لبرامج الأسلحة أو تحليل الأنماط في بيانات السواتل.

الذكاء الاصطناعي المتصل بالتحقق من تحديد الأسلحة وصنع القرار في حالات الأزمات

يمكن أن يساعد الذكاء الاصطناعي في التحقق من اتفاقات تحديد الأسلحة. ويرجع ذلك إلى قدرته على تحليل كميات كبيرة من البيانات – كتلك المستقاة من مصادر من قبيل صور السواتل – وتصنيف الأجسام المختلفة. إن هذه القدرة تجعل من الممكن تحديد الأعتدة العسكرية، مثل الدبابات والصواريخ والثكنات العسكرية، أو الأنشطة العسكرية، مثل تحركات القوات والتدريبات. وإضافة إلى ذلك، وكما سبق ذكره، يمكن اكتشاف برامج الأسلحة غير المشروعة بسهولة أكبر بواسطة الذكاء الاصطناعي. وبالتالي سيكون من الأصعب بكثير ارتكاب انتهاكات لاتفاقات تحديد الأسلحة ويمكن للدول الأطراف أن تتأكد من التزام الجميع بأحكام الاتفاقات.

إن الحصول على قدر أكبر وأفضل من المعلومات المستندة إلى قدرة الذكاء الاصطناعي على تحليل وتصنيف بيانات أجهزة الاستشعار لا يمكنها أن تسهل تنفيذ اتفاقات تحديد الأسلحة فحسب، بل يمكنها أيضا أن تسهم في تحسين عملية صنع القرار في الحالات التي تكون فيها حالات التوتر العسكري بين الدول مرتفعة بشكل خاص. ويمكن أن يستفيد القادة السياسيون والعسكريون من الإلمام بالحالة المدعوم بالذكاء الاصطناعي لخفض التصعيد في الأزمات.

السلام والأمن وميثاق الأمم المتحدة

ثمة تحدٍ خاص مرتبط بتطبيقات الذكاء الاصطناعي في المجال العسكري يتمثل في المخاطر المحتملة على السلام والأمن من خلال التصعيد غير المقصود وسوء الفهم الناتج عن استخدام الذكاء الاصطناعي. واستخدام التعلّم الآلي يضيف طبقة إضافية من التعقيد لأن طريقة عمل النظام قد لا تكون مفهومة تماماً من قبل جميع الجهات المعنية.

كما أن من الضروري اتخاذ تدابير ووضع ضوابط تضمن المساءلة والمسؤولية وتخفف من التحيز الخوارزمي تمهيدا لاستخدام الذكاء الاصطناعي في نظم دعم القرار في ما يتعلق بالتفاعل بين الإنسان والآلة وضرورة إمساك البشر بزمام المسؤولية.

ويتعين التخفيف من كل هذه المخاطر من خلال الرقابة والتدابير التي تأخذ في الاعتبار التحديات المحددة التي تحملها معها هذه التكنولوجيات.

ويشـــار إلى أن المادة 36 من البروتوكول الإضـــافي إلى اتفاقيات جنيف المعقودة في 12 آب/أغســـطس 1949 والمتعلق بحماية ضـــحايا النزاعات الدولية المســـلحة (البروتوكول الأول) تفرض الالتزام باستعراض مشروعية كل الأسلحة أو الوسائل أو الأساليب الحربية الجديدة قبل استخدامها في نزاع مسلح.

ويمكن أيضـــا تطبيق الذكاء الاصـــطناعي بطريقة تدعم التنفيذ الفعال للالتزامات بموجب القانون الدولي الإنساني، لا سيما عندما يتعلق الأمر بحماية المدنيين، باعتباره التزاما إيجابيا وعملا إيجابيا، بما في ذلك من خلال المشاريع والبحوث والتطبيقات المصممة خصيصا لهذه المهمة.

أطر للتعاون المتعدد الأطراف وتبادل المعلومات

بالنظر إلى التطور السـريع لمسـألة الذكاء الاصـطناعي في المجال العسـكري وإلى ما تطرحه من تحديات لجميع الدول، تتسـم المناقشـات المتعددة الأطراف وصـيغ تبادل الخبرات وأفضـل الممارسـات بقدر كبير من الأهمية. وفي هذا الصـدد، أيدت النمسـا الإعلان السـياسـي بشـأن الاسـتخدام العسـكري المسـؤول للذكاء الاصـطناعي وخاصـية التشـغيل الذاتي. وتعمل النمسـا، بصـفتها رئيسـا مشـاركا، إلى جانب ألمانيا، للفريق العامل المعني بالإشراف على هذا الإعلان على تيسير تبادل أفضل الممارسات في مواجهة التحديات وصـياغة السـياسـات في هذا المجال. وأيدت النمسـا أيضـا مخطط العمل الصـادر عن قمة تسـخير الذكاء الاصـطناعي المسـؤول في المجال العسـكري، وكذلك إعلان باريس بشـأن الحفاظ على التحكم البشـري في منظومات الأسلحة التي تعتمد على الذكاء الاصطناعي.

العلاقة بين عمل المجتمع الدولي المتعلق بالذكاء الاصـطناعي في المجال العسـكري وعمله في مجال منظومات الأسلحة الذاتية التشغيل

ضـمن النطاق الأوسـع لتطبيق الذكاء الاصـطناعي والتشـغيل الذاتي في المجال العسـكري، ثمة مسـألة محددة تتعلق بمنظومات الأسـلحة الذاتية التشـغيل يتعين تسـليط الضـوء عليها. وتثير منظومات الأسـلحة الذاتية التشـغيل هواجس خاصـة من منظور قانوني وأخلاقي وأمني. وهذه المسـألة ليسـت محور قرار الجمعية العامة 239/79 إذ إن المناقشـات في إطار الأمم المتحدة مسـتمرة بالفعل منذ عام 2013، حيث أعربت غالبية متزايدة من الدول عن رغبتها في وضـع قواعد لمنظومات الأسـلحة الذاتية التشـغيل وفرض قيود عليها على الصـعيد الدولي. ولغرض هذا التقرير، ستقصـر النمسا تعليقاتها بغية التأكيد على موقفها المؤيد لوضـع صـك ملزم قانوناً بشأن منظومات الأسلحة الذاتية التشغيل، وستشير هنا إلى العمل الهام الذي يضـطلع به حالياً فريق الخبراء الحكوميين في إطار اتفاقية حظر أو تقييد اسـتعمال أسـلحة تقليدية معينة يمكن اعتبارها مفرطة الضرر أو عشوائية الأثر، وكذلك إلى الجهود التكميلية المبذولة في إطار أول قرار للجمعية العامة بشأن منظومات الأســـلحة الفتاكة الذاتية التشـــغيل (القرار 241/78)، الذي صـــدر بموجبه تقرير الأمين العام (A/79/88)، وقرار المتابعة (القرار 62/79)، الذي مهّد لتنظيم مشاورات غير رسمية بشأن منظومات الأسلحة الفتاكة الذاتية التشغيل من المقرر إجراؤها في نيويورك يومي 12 و 13 أيار/مايو 2025.

الاعتبارات المتصلة بالأطر القانونية للذكاء الاصطناعي

ينشـئ قانون الاتحاد الأوروبي للذكاء الاصـطناعي إطارا تشـريعيا للاتحاد الأوروبي يتعلق بنظُم الذكاء الاصطناعي في قطاعات مختلفة، ويهدف إلى تعزيز الثقة في تطبيقات الذكاء الاصطناعي وتسخير فوائد الذكاء الاصطناعي مع حماية حقوق الإنسان والحريات الأساسية والقيم الديمقراطية. ويؤكد على أهمية الشفافية والمساءلة والرقابة البشرية في تطوير ونشر منظومات الذكاء الاصطناعي مع تعزيز اليقين القانوني والابتكار والتنافسية. ولا يسري القانون على منظومات الذكاء الاصطناعي المطوَّرة لغرض الأنشطة العسكرية أو الدفاعية أو لأنشـطة الأمن القومي. بيد أن قانون الذكاء الاصـطناعي يطبق نهجا قائما على المخاطر، وهو ما قد يكون مفيدا عند التعامل مع مجموعة واسـعة من تطبيقات الذكاء الاصـطناعي المحتملة في المجال العسكري.

سبل المضي قدما

تثمِّن النمسـا العمل الذي جرى الاضـطلاع به في الأشـكال والمنتديات المختلف المذكورة في مساهمتها في ما يتعلق بتطبيقات الذكاء الاصطناعي في المجال العسكري، وهي على ثقة بأنه سيسهم في مجموعة من القواعد والمعايير الناشئة المتفق عليها دوليا لضمان الاستخدام المسؤول للذكاء الاصطناعي في المجال العسكري وفقًا للالتزامات القانونية الدولية والمبادئ الأخلاقية.

شيلي

[الأصل: بالإسبانية]

[11 نيسان/أبريل 2025]

ذكرت شيلي في وقت سابق أن التطوير السريع للتكنولوجيات الجديدة والناشئة مسألة هامة فيما يتعلق بالأمن الدولي ويطرح تحدياً لجميع البلدان. وقد تُنتِج هذه التكنولوجيات الجديدة، ولا سـيما الذكاء الاصـطناعي، فوائد هائلة لتنمية المجتمعات ورفاهها، لكنها في الوقت نفسـه تطرح تسـاؤلات كبيرة حول تداعيات اسـتخدامها في مجال الأمن والدفاع. ويمكن للتكنولوجيات الجديدة أن تولد فوائد هامة، لكنها يمكن أن تولد أيضاً مخاطر وصعوبات.

وفي هذا الصدد، ترى شيلي أنه سيكون من المستحسن إيجاد تفاهم مشترك بشأن استخدام مسؤول للذكاء الاصطناعي في المجال العسكري والأمني وبشأن تطوير واستخدام ما يسمى منظومات الأسلحة الفتاكة الذاتية التشغيل. وتدعم شيلي الجهود المتعددة الأطراف الرامية إلى استحداث وتعزيز منتديات للحوار والمناقشة بين البلدان، بهدف إيجاد مجالات للتفاهم المتبادل وتوافق الآراء حول استخدام هذه التكنولوجيات الجديدة.

وتضطلع شيلي بدور طليعي في مجال الذكاء الاصطناعي نظراً إلى التقدم الكبير الذي أحرزه البلد في تهيئة الظروف المؤاتية لنشـر هذه التكنولوجيا، فضـلاً عن التقدم الرائد الذي أحرزه في ما يتعلق بالسـياسـات والمناقشـات التنظيمية بشأن الذكاء الاصطناعي. وفي تشرين الأول/أكتوبر 2021، أطلقت شيلي سياستها الوطنية الأولى بشأن الذكاء الاصـطناعي التي وُضـعت بالتعاون مع جهات فاعلة مختلفة من القطاعين العام والخاص. وتركز هذه السياسة على ثلاث ركائز أساسية هي: عوامل التمكين؛ واستخدام التكنولوجيا وتطويرها؛ واستحداث الأطر الناظمة والأخلاقية لضمان الاستخدام المسؤول والآمن للذكاء الاصطناعي.

وفي عام 2024، أصدرت شيلي نسخة محدَّثة من سياستها الوطنية بشأن الذكاء الاصطناعي تتضمن مواضيع فرعية جديدة مثل التنسيق الدولي، والبيئة وأزمة المناخ، والشمولية وعدم التمييز، والأطفال والمراهقين، والثقافة والحفاظ على التراث. وتُستكمل هذه السياسة بخطة عمل تتضمن أكثر من 100 تدبير من المقرر استكمالها بحلول عام 2026، تتناول مجالات من قبيل التعليم والصحة والبيئة والثقافة. وأُدرجت أيضا المبادئ المنصوص عليها في التوصية الخاصة بأخلاقيات الذكاء الاصطناعي الصادرة عن منظمة الأمم المتحدة للتربية والعلم والثقافة (اليونسكو) في السياسة الوطنية الجديدة بشأن الذكاء الاصطناعي، بحيث تكون هذه السياسة منسجمة مع أحدث الأطر الدولية.

وكانت شيلي أول بلد في العالم يطبق منهجية تقييم الاستعداد، وهي أداة طورتها اليونسكو لتحديد مدى جاهزية كل بلد لتطبيق الذكاء الاصطناعي بطريقة أخلاقية ومسؤولة. وبذلك أكدت شيلي من جديد التزامها بتنفيذ توصية اليونسكو الخاصة بأخلاقيات الذكاء الاصطناعي في قواعدها التنظيمية الوطنية. وتشجع شيلي على التطوير الأخلاقي والمسؤول لهذه التكنولوجيا، كما يتجلى ذلك من مشاركة البلد في مؤتمرات القمة حول الذكاء الاصطناعي التي نظمتها المملكة المتحدة (2023) وجمهورية كوريا (2024) وفرنسا (2025).

وفي ما يتعلق بالتشريعات، تجرى حاليا في شيلي مناقشة مشروع قانون لوضع قواعد تنظيمية لمنظومات الذكاء الاصطناعي باستخدام نهج قائم على المخاطر بهدف الترويج لتطوير وتنفيذ هذه المنظومات مع دعم المبادئ الديمقراطية والحقوق الأساسية للأفراد.

وفي مجال الدفاع والأمن، دعمت شيلي وشاركت بنشاط في مؤتمري القمة لتسخير الذكاء الاصطناعي المسؤول في المجال العسكري، اللذين عُقدا في لاهاي (2023) وسول (2024). وأيدت شيلي الوثائق المعتمدة في مؤتمري القمة (الدعوة إلى العمل (2023) ومخطط العمل (2024)). وقدمت شيلي أيضا الدعم لعمل الهيئة العالمية لتسخير الذكاء الاصطناعي المسؤول في المجال العسكري.

وفي 13 و 14 حزيران/يونيه 2024، عُقدت في شيلي حلقة عمل إقليمية حول الاستخدام المسؤول للذكاء الاصطناعي في المجال العسكري والمجال الأمني الأوسع نطاقاً. ونظمت وزارة الخارجية في شيلي ووزارة الخارجية في كوستاريكا حلقة العمل برعاية مملكة هولندا وجمهورية كوريا. وقدم كذلك مركز دراسة القانون والتكنولوجيا والمجتمع في جامعة شيلي ومركز الحوار في المجالات الإنسانية ومقره جنيف الدعمَ والمساعدة في تنظيم هذه المناسبة. وشارك في حلقة العمل ممثلون عن الأرجنتين، وأوروغواي، وباراغواي، والبرازيل، وترينيداد وتوباغو، وجامايكا، والجمهورية الدومينيكية، والسلفادور، وشيلي، وكوستاريكا، وكولومبيا، والمكسيك، فضلا عن جمهورية كوريا ومملكة هولندا. وشارك أيضا ممثلون عن وزارة الدفاع الوطني والقوات المسلحة باسم شيلي.

وتعتقد شيلي أن تطبيقات الذكاء الاصطناعي في المجال العسكري والأمني يمكن أن تحقق فرصا وفوائد، مثل تعزيز عملية صنع القرار والتحليل الاستراتيجي، وعمليات لوجستية أكثر كفاءة، وتعزيز القدرات في مجال الدفاع السيبراني والأمن السيبراني – وبالتالي تعزيز أمن البنية التحتية الحيوية – فضلا عن تسهيل التخطيط لبعثات حفظ السلام والمساعدة الإنسانية المعقدة. ويمكن لتطبيقات الذكاء الاصطناعي أيضًا أن تحسّن القدرات على التحقق والرصد لتحديد الأسلحة وإنفاذ أنظمة مراقبة الأسلحة.

وتعتقد شيلي أنه يجب تطوير تكنولوجيات الذكاء الاصطناعي ونشرها واستخدامها وفقا للقانون الدولي، بما في ذلك، حسب الاقتضاء، ميثاق الأمم المتحدة والقانون الدولي الإنساني والقانون الدولي لحقوق الإنسان وغيرها من الأطر القانونية ذات الصلة.

وترى شيلي أن من المهم وضع تدابير رقابية وأمنية لمنع الجهات الفاعلة غير المسؤولة من الحصول على قدرات الذكاء الاصطناعي التي قد تكون ضارة في المجال العسكري وإساءة استخدامها، بما في ذلك المنظومات التي تستخدم الذكاء الاصطناعي، مع الأخذ في الاعتبار أن أي تدابير من هذا القبيل ينبغي ألا تقوض الوصول المتكافئ إلى فوائد قدرات الذكاء الاصطناعي في المجالات الأخرى غير العسكرية.

وعلى غرار ذلك، تعتقد شيلي أن من المهم أن تتضافر الجهود لمنع استخدام تكنولوجيات الذكاء الاصطناعي لتسهيل انتشار أسلحة الدمار الشامل بين الدول والجهات الفاعلة من غير الدول، بما في ذلك الجماعات الإرهابية، والتأكيد على ضرورة استخدام تكنولوجيات الذكاء الاصطناعي لدعم جهود نزع السلاح وتحديد الأسلحة وعدم الانتشار، بدلا من إعاقتها. ومن الأهمية بشكل خاص الحفاظ على التحكم البشري والمشاركة البشرية في جميع الإجراءات الحاسمة في توجيه وتنفيذ القرارات السيادية المتعلقة باستخدام الأسلحة النووية، دون المساس بالهدف النهائي المتمثل في إيجاد عالم خالٍ من الأسلحة النووية.

وتدعو شيلي إلى وضع تدابير لبناء الثقة، مثل تبادل المعلومات والمشاورات بين الدول بشأن الممارسات الجيدة والدروس المستفادة. وفي هذا الصدد، ترى شيلي أن من المهم بالنسبة للبلدان أن تضع وتؤسس استراتيجيات ومبادئ ومعايير وسياسات وأطر وقوانين وطنية لكفالة الاستخدام المسؤول للذكاء الاصطناعي في المجال العسكري. ويمكن لتدابير بناء الثقة أن تكون أداة فعالة لتطوير آليات الاحتواء والرقابة والمصداقية على المستويين الوطني والدولي، وبالتالي تعزيز الشفافية.

وبالمثل، تعتقد شيلي أنه من الضروري تقليص الفجوات الرقمية والمتعلقة بالذكاء الاصطناعي بين البلدان المتقدمة والبلدان النامية، وترى أنه من الضروري تعزيز الفهم والوعي بآثار الذكاء الاصطناعي في المجال العسكري، بما في ذلك تبادل المعارف ومشاركة الممارسات الجيدة والدروس المستفادة بين جميع الدول.

وفي هذا الصدد، تعتقد شيلي أنه من الضروري وضع مبادرات وبرامج تهدف إلى تعزيز بناء القدرات، لا سيما في البلدان النامية، لتعزيز مشاركتها الكاملة في المناقشات حول إدارة الذكاء الاصطناعي في المجال العسكري. وتدرك شيلي أن بناء القدرات يمكن أن يساعد البلدان أيضًا على اكتساب فهم أعمق للذكاء الاصطناعي في المجال العسكري وتيسير التطوير المسؤول والقانوني لقدرات الذكاء الاصطناعي العسكرية ونشرها واستخدامها. وسيمكّن بناء القدرات أيضا من تجهيز البلدان للمشاركة بفعالية أكبر في الحوار والمناقشات على الصعيد الدولي.

وترى شيلي أن من المهم تعزيز التعاون الدولي لبناء القدرات، وتشجيع الحوار والنقاش على المستويات الوطني والإقليمي ودون الإقليمي والأقاليمي، بما في ذلك من خلال برامج التدريب والمؤتمرات وحلقات العمل والحلقات الدراسية للمسؤولين الدبلوماسيين والسياسيين والتقنيين، بهدف سد الفجوة المعرفية المتعلقة بتطوير ونشر واستخدام الذكاء الاصطناعي في المجال العسكري على نحو مسؤول.

وتقدّر شيلي أهمية تعزيز النقاش والحوار الإقليمي ودون الإقليمي حول الذكاء الاصطناعي في المجال العسكري وتعتبره أمرا مهما. ومن الجهود البارزة التي بُذلت لتحقيق هذه الغاية المؤتمر السادس عشر لوزراء الدفاع في الأمريكتين، الذي عُقد في الأرجنتين في الفترة من 13 إلى 16 تشرين الأول/ أكتوبر 2024، ولا سيما الوثيقة الختامية للمؤتمر، المعروفة باسم إعلان مندوزا، والتي تتضمن التوصيات التالية: تعزيز الاستخدام الأخلاقي للذكاء الاصطناعي في الدفاع؛ ومراعاة التنوع الاقتصادي والتكنولوجي للدول الأعضاء في المؤتمر؛ ودعم آليات تعزيز الثقة المتبادلة والتعاون الإقليمي على نطاق نصف الكرة الأرضية بحيث تتيح تلك الآليات للدول الأعضاء في المؤتمر تبادل المعارف

والممارسات الجيدة، ووضع معايير قائمة على توافق الآراء، وبناء القدرات التكنولوجية لتطبيق الذكاء الاصطناعي في مجال الدفاع.

وأخيراً، تعتقد شيلي أن مشاركة جميع الأطراف المعنية، بما في ذلك المجتمع المدني والأوساط الأكاديمية والصناعية والتقنية والقطاع الخاص والمنظمات الإقليمية والدولية، في المناقشات والحوار حول استخدام الذكاء الاصطناعي في المجال العسكري أمر أساسي.

الصين

[الأصل: بالصينية]
[11 نيسان/أبريل 2025]

إن ما يشهده الذكاء الاصطناعي في المجال العسكري من تطور سريع وتطبيق واسع النطاق يعيد تشكيل نماذج حروب المستقبل ويطرح في الوقت نفسه تحديات محتملة أمام السلام والأمن الدوليين. ففي الوقت الذي يواجه فيه العالم تحديات متعددة للسلام والأمن، ينبغي لجميع الأطراف أن تسعى إلى التوصل إلى توافق في الآراء، من خلال الحوار والتعاون، على تنظيم التطبيقات العسكرية للذكاء الاصطناعي، وتشجيع تطوير إطار حوكمة مفتوح ومنصف وفعال لأمن الذكاء الاصطناعي، وتقليل المخاطر لضمان أن تظل تكنولوجيات الذكاء الاصطناعي آمنة وموثوقة ويمكن التحكم فيها وضمان تطورها دائمًا بما يعود بالنفع على تقدم ركب الحضارة الإنسانية.

وما فتئت الصين تشارك بطريقة مسؤولة وبناءة في الحوكمة العالمية للتطبيقات العسكرية للذكاء الاصطناعي. وإننا ندعو إلى التمسك بمفهوم "نهج محوره الإنسان في التطبيقات العسكرية للذكاء الاصطناعي" والتمسك برؤية قائمة على الأمن المشترك والشامل والتعاوني والمستدام، سعيا لبناء مجتمع ذي مستقبل مشترك للبشرية. وفي عام 2021، قدمت الصين ورقة موقف، في إطار الاتفاقية المتعلقة بأسلحة تقليدية معينة، عن تنظيم التطبيقات العسكرية للذكاء الاصطناعي، اقترحت فيها آراء وتوصيات منهجية بشأن التطوير المسؤول للذكاء الاصطناعي واستخدامه في المجال العسكري تتعلق بالأمن الاستراتيجي، والسياسات العسكرية، والقانون والأخلاقيات، والأمن التكنولوجي، وعمليات البحث والتطوير، وإدارة المخاطر والرقابة، ووضع القواعد والتعاون الدولي. وفي عام 2023، اقترحت الصين مبادرة الحوكمة العالمية للذكاء الاصطناعي، داعيةً جميع البلدان، ولا سيما الدول الكبرى، إلى أن تتخذ موقفا حصيفا ومسؤولا إزاء البحث في تكنولوجيات الذكاء الاصطناعي في الميدان العسكري، وتطويرها، وتطبيقها. وتشمل مقترحاتنا المحددة ما يلي:

أولاً، ينبغي اتباع نهج حصيف ومسؤول. وينبغي لجميع الدول، ولا سيما الدول الكبرى، أن تمتنع، أثناء تطويرها لقدراتها الدفاعية الوطنية المشروعة، عن السعي إلى تحقيق التفوق العسكري المطلق من خلال الذكاء الاصطناعي وعن تقويض المصالح الأمنية المشروعة للآخرين. وينبغي بذل الجهود لتجنب حالات سوء التفاهم والحسابات الخاطئة ومنع حدوث سباق تسلح في هذا المجال.

ثانياً، ينبغي التمسك بنهج محوره الإنسان. فمن الضروري أن نعتبر دائماً أن الإنسان هو من يتحمل المسؤولية في نهاية المطاف وأن نضمن كون منظومات الأسلحة ذات الصلة تحت تحكم البشر.

وينبغي للتطبيقات العسكرية للذكاء الاصطناعي أن تحترم وتحمي الكرامة الإنسانية وحقوق الإنسان وتحترم القيم الإنسانية المشتركة.

ثالثًا، ينبغي مراعاة المبدأ الأساسي المتمثل في "تسخير الذكاء الاصطناعي لخدمة الخير". وينبغي لتطبيق الذكاء الاصطناعي في المجال العسكري أن يسهم في الحفاظ على السلام، والامتثال للقانون الدولي الإنساني وغيره من القوانين الدولية الواجبة التطبيق، وأن يسعى إلى الحد من الخسائر الجانبية.

رابعاً، ينبغي تطبيق الحوكمة المرنة. وينبغي لنا أن نقوم بتعزيز التقييم الاستشرافي للمخاطر وتدريب الموظفين على الذكاء الاصطناعي، واتخاذ التدابير اللازمة للتخفيف من المخاطر والحد من مخاطر الانتشار، دون أن نعيق الابتكار والاستخدامات السلمية للتكنولوجيات.

خامسًا، ينبغي التمسك بتعددية الأطراف. وينبغي أن ندعم الأمم المتحدة في الاضطلاع بدورها الواجب، وأن نرحب بإيجاد منصات شاملة للجميع لإجراء مناقشات فيها يشارك فيها جميع الأطراف، وأن نبذل جهودًا لإنشاء أطر للحوكمة تقوم على المشاركة العالمية والتوافق الواسع في الآراء.

وترى الصين أن أهمية الذكاء الاصطناعي في المجال العسكري ينبغي أن تُقيَّم بموضوعية. ومن الضروري توجيه تطوير الذكاء الاصطناعي العسكري في الاتجاه الصحيح مع منع النمو غير الخاضع للتنظيم في الوقت ذاته. وفي المرحلة المقبلة، ينبغي للمجتمع الدولي أن يتعاون على تحقيق أقصى قدر من الفوائد مع تقليل الأضرار في الوقت نفسه. وتقترح الصين الأفكار والاقتراحات التالية:

أولاً، وضع مبادئ توجيهية واضحة – يجب إيلاء الأمن والتنمية نفس القدر من الاهتمام. ولا بد من الالتزام بمقاصد ميثاق الأمم المتحدة ومبادئه، ومراعاة المعايير الأساسية التي تحكم العلاقات الدولية، وضمان عدم تحول تكنولوجيا الذكاء الاصطناعي إلى أداة لغزو البلدان الأخرى والسعي إلى الهيمنة. والصين على استعداد لإجراء مزيد من تبادل الآراء مع جميع الأطراف بشأن مفهوم "نهج محوره الإنسان في التطبيقات العسكرية للذكاء الاصطناعي" وبناء توافق في الآراء بشكل مستمر.

ثانيًا، تحسين تدابير الحوكمة – في سياق الحالة الراهنة التي بلغها تطوير الذكاء الاصطناعي وتطبيقه، ينبغي لنا أن نشجع إنشاء نظام للاختبار والتقييم، وأن نطبق حوكمة مرنة، وأن ننفذ إدارة متعددة المستويات وقائمة على الفئات من أجل استجابة سريعة وفعالة. وينبغي لجميع البلدان، استنادًا إلى ظروفها الوطنية، أن تقوم بإنشاء وتحسين نظمها القانونية والتنظيمية المحلية، وصقل المبادئ التوجيهية الأخلاقية ذات الصلة، وتعزيز التعليم والتدريب، بغية تعزيز سلامة تكنولوجيات الذكاء الاصطناعي وموثوقيتها وإمكانية التحكم فيها.

ثالثاً، تعزيز التعاون الدولي – ينبغي لجميع البلدان الالتزام بمبدأي الانفتاح وشمول الجميع، والانخراط في الحوار وتبادل الآراء لتعزيز التفاهم المتبادل، وتعزيز تنسيق السياسات والتعاون في مجال بناء القدرات فيما يتعلق بحوكمة الذكاء الاصطناعي من أجل تحسين مستوى الحوكمة باستمرار.

مصر

[الأصل: بالإنكليزية]
[11 نيسان/أبريل 2025]

عملاً بقرار الجمعية العامة 239/79، تود حكومة جمهورية مصر العربية أن تعرض آراءها بشأن ما يتيحه تطبيقُ الذكاء الاصطناعي في المجال العسكري من فرص ويطرحه من تحديات بالنسبة إلى السلام والأمن الدوليين.

ويمثل قرار الجمعية العامة 239/79 خطوة هامة لتعزيز تعددية الأطراف بشأن موضوع الذكاء الاصطناعي في المجال العسكري ونحو رفع مرتبته على جدول الأعمال السياسي. ويأتي ذلك في أعقاب دعوة الأمين العام إلى وضع معايير وقواعد ومبادئ حول تصميم وتطوير واستخدام التطبيقات العسكرية للذكاء الاصطناعي بمشاركة جميع الجهات المعنية.

ومع التسليم بأن القرار المذكور أعلاه، والذي تُعرض بموجبه هذه الآراء، يهدف إلى التركيز بشكل خاص على مجالات أخرى غير منظومات الأسلحة الفتاكة الذاتية التشغيل، فمن الأهمية بمكان التأكيد مجدداً على موقف مصر الثابت بأن أي مناقشة مجدية لهذا الموضوع لا يمكن أن تُغفِل أولويةَ تناوُل كل الأبعاد الأخلاقية والقانونية والأمنية المحيطة بمنظومات الأسلحة الفتاكة الذاتية التشغيل التي تمثل التهديد الأكثر إلحاحاً لصون السلام والأمن الدوليين في ما يتعلق بالتطبيقات العسكرية للذكاء الاصطناعي.

إن الاتفاق على حظر ملزم قانوناً لمنظومات الأسلحة الفتاكة الذاتية التشغيل التي تعمل دون تحكّم أو مراقبة بشريين والتي لا يمكن استخدامها بما يتوافق مع القانون الدولي الإنساني، كما اقترح الأمين العام سابقاً، هو أكثر مسارات العمل فعالية وواقعية. واتّباعُ نهج يتضمن مستويين من الحظر والتقييد و/أو التنظيم – يشمل حظر منظومات الأسلحة التي تعمل دون تحكّم بشري والتنظيم المتعلق بالمنظومات الأخرى – ضروري لإنشاء الهيكل القانوني العالمي اللازم الذي من شأنه أن يوفر بيئة مؤاتية لتحقيق أقصى قدر من فوائد الفرص الجديدة التي توفرها التطبيقات العسكرية للذكاء الاصطناعي مع مواجهة التحديات ذات الصلة بطريقة واقعية وفعالة وفي الوقت المناسب.

ولا يزال مشهد السياسة العامة الدولية المحيطة بالذكاء الاصطناعي في المجال العسكري أبعد من أن يكون موحدا. إن مصر تتابع عن كثب المبادرات الدولية المتعددة في هذا الشأن، والتي تُظهر وعيا متزايدا بالمخاطر المرتبطة بها. ومع ذلك، كشفت المداولات خلال هذه المبادرات عن وجود اختلافات في الآراء وتصورات التهديد والأولويات، وبناءً على ذلك، يتعين علينا توخّي الحذر من مغبّة خلق إطار سياساتي مجزأ أو آليات متنافسة، كما هو الحال في مجالات أخرى للتكنولوجيات الجديدة والناشئة.

وثمة حاجة واضحة إلى تبسيط هذه المبادرات وإلى وضعها تحت مظلة الأمم المتحدة بما يضمن شموليتها وفعاليتها. وتمثل الأمم المتحدة وآليتها لنزع السلاح المنبر الوحيد الفعال والشامل لوضع القواعد الدولية اللازمة والإطار المعياري اللازم، سيّما وأن الوتيرة السريعة للتطورات التكنولوجية لا تزال تفوق بكثير التنظيم اللازم على الصعيد الدولي.

لذا، من الضروري استحداث منصة عالمية ومستقلة وموثوقة وذات مسار واحد تحت رعاية الأمم المتحدة لمناقشة كيفية حوكمة الذكاء الاصطناعي في المجال العسكري في المستقبل. وستكون العملية

المتوخاة التي تقودها الأمم المتحدة مصمَّمة خصيصا لتلافي ثنائيات ناشئة معينة تؤدي إلى نتائج عكسية. ومن هذه الثنائيات ثنائية الجهود المشروعة لضمان الامتثال القانوني والأخلاقيات مقابل الميل إلى تحقيق المصالح العسكرية دون إيلاء الاعتبار الواجب للآثار الإنسانية.

وسيجرى التأكيد كذلك على أنه، مع تقديرنا للمناقشات التي أُجريت في إطار فريق الخبراء الحكوميين المعني بمنظومات الأسلحة الفتاكة الذاتية التشغيل تحت مظلة اتفاقية حظر أو تقييد استعمال أسلحة تقليدية معينة يمكن اعتبارها مفرطة الضرر أو عشوائية الأثر، لا يمكن أن يكون هذا المنبر بديلاً من عملية الأمم المتحدة المتوخاة بشأن تطبيقات الذكاء الاصطناعي في المجال العسكري، إذ إن فريق الخبراء الحكوميين ليس عالمياً بطبيعته وهو غير مكلَّف بتناوُل موضوع متنوع ومتفرّع كهذا. ومن المؤسف أيضاً أن التقدم المحرز في إطار فريق الخبراء الحكوميين لا يزال طفيفا إلى حد كبير ولم تحقَّق نتائج محددة في هذا الصدد حتى الآن.

وبالاقتران مع الفرص التي تتيحها تقنيات الذكاء الاصطناعي، تكمن مجموعة من المخاطر الملازمة لخصائص هذه التكنولوجيات، والتي يمكن أن تعمل بشكل غير متوقع وغير قابل للتفسير. وتشمل هذه المخاطر المعلومات المضللة والتصعيد غير المقصود والمخاطر السيبرانية، فضلا عن إساءة الاستخدام والانتشار من قبل الجهات غير التابعة للدول. ويمكن أن تكون المخاطر مستجدة أو يمكن أن تجعل المخاطر القائمة أكثر تعقيداً.

ومن المسلَّم به على نطاق واسع أن هناك مجموعة واسعة من التطبيقات العسكرية الممكنة للذكاء الاصطناعي. ومع ذلك، فإن الجهود الجادة التي تُبذل لبلورة كيفية حوكمتها في المستقبل ستحدد الترتيب الصحيح للأولويات من حيث المخاطر الكامنة فيها وتأثيرها على السلام والأمن. ويهدف ذلك إلى ضمان إجراء مناقشات مركَّزة ومنظَّمة، مع تجنب التشتت غير المبرر. ومن هذا المنطلق، ترى مصر بقوة أن التركيز يجب أن ينصبّ، إلى جانب مسألة منظومات الأسلحة الفتاكة الذاتية التشغيل، على قدرات المنظومات الذاتية التشغيل أو شبه ذاتية التشغيل التي تتيح استخدام القوة و/أو تخفض عتبة استخدام القوة، والتي قد تؤدي تالي إلى ديناميات سباق تسلح إضافية تشمل الأسلحة التقليدية وغير التقليدية. إن احتمال حصول تشغيل ذاتي متزايد للأسلحة النووية والأسلحة التقليدية المتطورة، مثل الصواريخ الفرط صوتية، من شأنه أن يخلق مخاطر غير معروفة وأن يحوّل مستقبل النزاع بطرق لا يمكن التنبؤ بها.

وينبغي التركيز أيضاً على أنشطة القيادة والتحكم واختيار الأهداف عوض التركيز على التخطيط اللوجستي والاستخبارات، والمراقبة والاستطلاع، نظراً لتأثيراتها الأقل تعطيلاً. وعلى غرار ذلك، يتعين التركيز بشكل أكبر على القدرات الهجومية عوض القدرات الدفاعية.

وينبغي للمداولات المتوخاة في إطار العملية المنشودة التي تقودها الأمم المتحدة أن تهدف أولاً إلى التوصل إلى تفاهم حول العناصر الرئيسية التي يقوم عليها تطوير الذكاء الاصطناعي ونشره واستخدامه في المجال العسكري. وتشمل هذه العناصر:

- الامتثال التام للقانون الدولي الواجب التطبيق، بما في ذلك المبادئ الأساسية للقانون الدولي الإنساني، مثل الضرورة والتناسب والتمييز، فضلاً عن الاعتبارات الأخلاقية الأخرى طوال دورة حياة ومراحل تطبيقات الذكاء الاصطناعي في المجال العسكري.

- مركزية الحفاظ على العنصر البشري طوال دورة حياة التطبيقات العسكرية للذكاء الاصطناعي، بما في ذلك الحكم والتدخل والرقابة والتحكم من قبل البشر، باعتبارها عوامل تمكين رئيسية للحفاظ على المساءلة. ومن الضروري ضمان أن تظل كل البرامجيات والخوارزميات والتصاميم التي تنطوي على استخدام تطبيقات الذكاء الاصطناعي في المجال العسكري خاضعة للمراجعة البشرية البالغة الأهمية ومبدأ قابلية التفسير. ومع أن الحكومات تزعم أن التحكم البشري بالمنظومات القائمة على الذكاء الاصطناعي يتم الحفاظ عليه من منطلق عقائدي، فقد يميل البعض إلى جعل منظومات أسلحتهم ذاتية التشغيل بقدر أكبر تحقيقا لمصالحهم العسكرية.

- التوازن بين التخفيف من مخاطر الانتشار على الجهات غير التابعة للدول وكبح الاستخدام الخبيث مقابل الحفاظ على حقوق الدول في الحصول على الذكاء الاصطناعي والتكنولوجيات ذات الاستخدام المزدوج. ومن الأهمية بمكان تجنُّب اعتماد أية آليات مراقبة دولية بشكل اعتباطي أو فرض أي نوع من الرقابة التمييزية على الصادرات.

- عنصر لبناء القدرات بهدف ضمان الاستثمار المناسب في رأس المال البشري ونقل التكنولوجيا وتبادل المعارف وأفضل الممارسات على نحوٍ يحفظ حق البلدان النامية في الاستفادة من الفوائد المحتملة للتطبيقات العسكرية المختلفة للذكاء الاصطناعي وبهدف سد الفجوة الرقمية.

- حدود الذكاء الاصطناعي في المجال العسكري وتفاعله مع التكنولوجيات الجديدة والناشئة الأخرى. ومن الملائم مناقشة سبل ضمان التكامل مع العمليات الأخرى التي تقودها الأمم المتحدة، بما في ذلك الفريق العامل المفتوح العضوية المعني بأمن تكنولوجيا المعلومات والاتصالات واستخدامها، وذلك نظراً، على سبيل المثال، إلى أوجه التقاطع بين الذكاء الاصطناعي والعمليات السيبرانية. وإضافة إلى ذلك، ستركز المناقشات بشكل أساسي على المجال العسكري بصرف النظر عن المجالات الأمنية الأوسع نطاقاً.

وأخيرًا، من المهم ضمان الشمولية والإنصاف في تحديد مسارات الحوكمة من أجل ذكاء اصطناعي مسؤول وخاضع للمساءلة ومتمحور حول الإنسان ضمن منظورات الأمم المتحدة للجهات المعنية المتعددة التي تقدم إسهامات رئيسية تغذي المناقشات المتعلقة بالسياسات. ومع ذلك، ينبغي لمشاركة تلك الجهات ألا تمسّ بالحق السيادي للدول في عملية صنع السياسات.

السلفادور

[الأصل: بالإسبانية]
[10 نيسان/أبريل 2025]

معلومات أساسية

كان لاستخدام الذكاء الاصطناعي في المجال العسكري في السنوات الأخيرة دور مهم للغاية. وقد ثبت في العديد من التقارير أن هذه التكنولوجيات الجديدة تزداد تطوراً وانتشاراً، الأمر الذي جعل من الممكن استخدام هذه الأدوات الحاسوبية في عمليات التخطيط وصنع القرار العسكري، بما في ذلك ما يتعلق منها بمن أو ما يستهدفه الهجوم. ويثير هذا الأمر العديد من التساؤلات حول الأثر العام والتبعات القانونية والمخاطر التي يتعرض لها المدنيون نتيجة استخدام هذه التكنولوجيات. ومن الأمثلة على ذلك النقاش الذي دار خلال مفاوضات متعددة الأطراف بشأن

التبعات السياسية والقانونية والإنسانية المترتبة على هذه التكنولوجيات فيما يتعلق بمنظومات الأسلحة ذاتية التشغيل. بيد أن نطاق التطبيقات العسكرية للذكاء الاصطناعي أوسع بكثير.

ولهذا السبب، نحن بحاجة إلى توسيع فهمنا لاستخدام الذكاء الاصطناعي وتطبيقاته في المجال العسكري، لا سيما فيما يتعلق بالمهام المحددة للاستهداف العسكري واستخدام القوة.

وقد اكتسبت مسألة الاستخدام المسؤول للذكاء الاصطناعي في المجال العسكري أهمية خاصة منذ المناقشات التي دارت في مؤتمر القمة الأول لتسخير الذكاء الاصطناعي المسؤول في المجال العسكري، الذي عقد في مملكة هولندا في شباط/فبراير 2023. وبدأت هذه المسألة أيضا تكتسب أهمية أكبر في اجتماعات فريق الخبراء الحكومي المعني بمنظومات الأسلحة الفتاكة الذاتية التشغيل، الذي يجتمع في جنيف.

ولئن كان من المهم الإشارة إلى أن تطبيقات الذكاء الاصطناعي واستخداماته قد نوقشت حتى الآن أساسا في سياق المناقشات حول منظومات الأسلحة الذاتية التشغيل، فإن استخدام الذكاء الاصطناعي في التطبيقات العسكرية هو قضية أوسع بكثير اتخذت أبعادًا جديدة ولا تنطبق فقط على التطبيقات التي تركز على التشغيل الذاتي لمنظومات الأسلحة ولكنها تنطبق أيضًا، وعلى وجه الخصوص، على التطبيقات التي تهدف إلى أتمتة بعض المهام العسكرية.

وبصورة عامة، فإن مناقشة الذكاء الاصطناعي هي قضية ناشئة لا تزال قيد البحث وتتطور بسرعة كبيرة، لدرجة أنه يجري تطوير مبادرات على المستويات الوطني والإقليمي والمتعدد الأطراف لمعالجة تأثيرها. ومن الواضح أن بلدان أمريكا اللاتينية ومنطقة البحر الكاريبي ليست في نفس مستوى البلدان المتقدمة من حيث التكنولوجيا وبناء القدرات التي تسهل تحديد وفهم الفرص والمخاطر الناجمة عن استخدام الذكاء الاصطناعي. ولذلك فمن المهم لهذه البلدان تكوين موقف وطني يؤهلها للمشاركة بفعالية في المناقشات الناشئة في المحافل والأوساط الدولية، وبذلك، تأمين التعاون لبناء القدرات والجوانب المحددة حتى تكون في صدارة هذه القضية وتفهم الفرص والمخاطر المحتملة المتعلقة بالأمن على المستويات الوطني والإقليمي والعالمي.

المبادرات التي شاركت فيها السلفادور

- شاركت السلفادور في قمة تسخير الذكاء الاصطناعي المسؤول في المجال العسكري التي عُقدت في مملكة هولندا عام 2023، وصادقت على الإعلان الصادر عن القمة (شباط/فبراير 2023).
- شاركت السلفادور في مؤتمر أمريكا اللاتينية ومنطقة البحر الكاريبي بشأن الأثر الاجتماعي والإنساني للأسلحة الذاتية التشغيل الذي اعتُمد فيه بيان بيلين (شباط/فبراير 2023).
- السلفادور هي جزء من "مجموعة الـ 16" في سياق مناقشات فريق الخبراء الحكومي المعني بمنظومات الأسلحة الفتاكة الذاتية التشغيل. وعلى الرغم من أن هذه المسألة تمثل محور تركيز متميزا، فهي تتعلق باستخدام الذكاء الاصطناعي في المجال العسكري.

الموقف الوطني

- قد يكون لبعض تطبيقات الذكاء الاصطناعي فوائد معينة في المجال العسكري، لا سيما تلك التطبيقات التي لا ترتبط بمهام تحديد الأهداف العسكرية والتعرف عليها أو التي لا ترتبط صراحةً باستخدام القوة، وهو ما ينطوي على مخاطر على المدنيين. وتتعلق هذه التطبيقات بمهام أخرى إدارية، مثل تحليل البيانات والتعلم الآلي غير المرتبط بالتفاعل البشري في العمليات العسكرية.

- بيد أن إساءة استخدام هذه التطبيقات يمكن أن يكون لها آثار سلبية، لا سيما فيما يتعلق بحماية المدنيين والبنية التحتية المدنية، وهما فئتان تحظيان بحماية خاصة بموجب قواعد القانون الدولي، بما في ذلك القانون الدولي الإنساني والقانون الدولي لحقوق الإنسان.

- من المهم أن يكون هناك نهج قائم على المخاطر يمكن من خلاله تنظيم بعض مهام الذكاء الاصطناعي أو حظرها، لا سيما المهام التي توضع فيها حدود للتحكم البشري الكبير في استخدام القوة وتلك التي تستنسخ أشكال التحيز الخوارزمية الناشئة عن استخدام قواعد بيانات غير تمثيلية أو التي تحتوي على بيانات تاريخية. وتشكل هذه المهام مخاطر تتعلق بحقوق الإنسان، وعلى المدى الطويل، بالأمن الدولي، خاصة عندما تُترك سلطة تقرير حياة الإنسان أو موته للآلة، أو عندما تحتوي هذه الأدوات على عناصر تكنولوجية متطورة للغاية، مثل التعلم الذاتي، التي قد تكون لها عواقب إنسانية أو اجتماعية أو اقتصادية أو سياسية أو حتى بيئية خطيرة.

- هناك حاجة ملحة في الوقت الحالي لإدخال قواعد تنظيمية مناسبة في مجال الذكاء الاصطناعي، فالقواعد التنظيمية ضرورية لضمان تطوير هذه التكنولوجيا بطريقة أخلاقية وآمنة. وسيساعد ذلك على حماية المستخدمين والمجتمع من التجاوزات والمخاطر المحتملة، كما سيشجع الابتكار من خلال توفير بيئة واضحة وآمنة للمطورين والباحثين.

- ولئن كان الهدف النهائي وضع صكوك قانونية ملزمة، فإن هذه التكنولوجيات تتقدم بمعدل يفوق تطور القانون الدولي في هذا المجال. ولهذا السبب، نعتقد أنه سيكون من المناسب اتباع نهج يركز على السلوك المسؤول، ويمكن استخدامه بعد ذلك لبناء أساس من الالتزامات القانونية الشاملة لمعالجة المشكلة بشكل أفضل.

- من المهم النظر في التحديات التي تطرحها التكنولوجيات الناشئة فيما يتعلق بالقضايا الأمنية. فعلى سبيل المثال، يمكن استخدام تكنولوجيات المواد، من قبيل الطباعة الثلاثية الأبعاد، لتصنيع الأسلحة الصغيرة والأسلحة الخفيفة؛ ويمكن استخدام التشغيل الآلي لتطوير روبوتات تتمتع بقدرات التشغيل الذاتي في المجال العسكري؛ وتعني الطبيعة المزدوجة لبعض استخدامات وتطبيقات الذكاء الاصطناعي أنها قد تكرر أشكال التحيز في مهام القيادة والتحكم في النزاعات المسلحة، مما يشكل مخاطر كبيرة على المدنيين.

- قد يؤدي فقدان التحكم أو توكيل الغير به في المجال العسكري إلى مخاطر غير مقصودة. ويمكن استخدام الذكاء الاصطناعي لزيادة القدرات البشرية، ولكن ضعف التحكم في السياق العسكري قد يشكل مخاطر أخرى يجب استكشافها بشكل كامل. ويجب استخدام الدعم المرتبط بالذكاء الاصطناعي في المجال العسكري لتعزيز أو إثراء عملية صنع القرار في سياقات محددة، لكنه يجب ألا يحل أبداً محل اتخاذ القرار والتفكير البشريين.

- يجب أن يتوافق استخدام الذكاء الاصطناعي في المجال العسكري مع القانون الدولي والقانون الدولي لحقوق الإنسان والقانون الدولي الإنساني ويجب أن يصب في الصالح العام.

- يجب أن تكون الدول قادرة على تعزيز قدراتها على تحديد المخاطر الناجمة عن إساءة استخدام الذكاء الاصطناعي وصلة ذلك بالقانون الدولي.

- ينبغي إشـــراك الجهات الفاعلة الأخرى المنخرطة في اســـتحداث وتطوير هذا النوع من التكنولوجيا، مثل الشركات الخاصة والأوساط الأكاديمية، في المناقشات المتعددة الأطراف، كما ينبغي تشجيع التعاون الدولي بين أصحاب المصلحة من أجل الاستفادة من المنافع التي توفرها الاستخدامات السلمية للذكاء الاصطناعي لدعم تنمية البلدان.

فنلندا

[الأصل: بالإنكليزية]
[11 نيسان/أبريل 2025]

يسر فنلندا أن تقدم آراءها بشأن قرار الجمعية العامة 239/79 المتعلق بالذكاء الاصطناعي في المجال العسكري وآثاره على الأمن الدولي، الذي اتُخذ في 24 كانون الأول/ديسمبر 2024، والذي تطلب فيه الجمعية إلى الأمين العام أن يلتمس آراء الدول الأعضاء بشأن "الفرص والتحديات التي يطرحها تطبيق الذكاء الاصطناعي في المجال العسكري على السلام والأمن الدوليين، مع التركيز بوجه خاص على مجالات غير منظومات الأسلحة الفتاكة الذاتية التشغيل".

ويشكل اعتماد مبادئ أو قواعد ناظمة دولية بشأن تطبيق الذكاء الاصطناعي في المجال العسكري عاملا أساسيا لضمان الامتثال للقانون الدولي، بهدف تعزيز الأمن والتقليل من احتمال نشوب النزاعات. وفي الوقت نفسه، من الضروري الإفساح في المجال أمام تطوير القدرات الدفاعية الوطنية الممتثلة للقانون الدولي. إن فنلندا ملتزمة بتطوير ونشر واستخدام قدرات الذكاء الاصطناعي في المجال العسكري بطريقة مسؤولة، وفقا للقانون الدولي، لا سيما القانون الدولي الإنساني، وعلى نحو لا يقوض السلام والأمن والاستقرار على الصعيد الدولي، مع مواصلة جهود البحث والتطوير والتجريب والابتكار في مجال تكنولوجيا الذكاء الاصطناعي.

لقد بات من المهم بشكل متزايد تحديد الآثار المترتبة على التكنولوجيات الثورية وتطوير وسائل لمواجهتها في السياسات الخارجية والأمنية والدفاعية. إن فنلندا تنشط في المشاركة في المناقشات العالمية حول تنظيم التكنولوجيا وذلك بدفاعها عن الحقوق الأساسية وحقوق الإنسان وتصديها للمخاطر ذات الصلة في تطوير وتطبيق الذكاء الاصطناعي وما يتصل بذلك من سياسات.

وإضافة إلى تحديد مخاطر التكنولوجيات الثورية، من المهم أيضا التعرف إلى الفرص التي توفرها هذه التكنولوجيات لمجالات من قبيل الأمن، وتطوير القدرات الدفاعية، والنمو الاقتصادي، والإنتاجية، والتنمية المستدامة، والكفاءة التكنولوجية، والاستثمارات القطاعية.

الفرص

تتيح التكنولوجيات الثورية فرصاً كبيرة للنهوض بقطاعات مختلفة ودفع عجلة التحول النظيف وتشجيع النمو الاقتصادي المستدام وتحسين الكفاءة والإنتاجية. كما أنها تنطوي على القدرة على تعزيز الأمن والتعليم والرفاه والصحة على النطاق العالمي.

ويوفر الذكاء الاصطناعي والتقنيات الناشئة الأخرى فرصا لتطوير القدرات الدفاعية ترسم ملامح مستقبل ساحات القتال ووسائل وأساليب الحرب بشكل أساسي. وتتيح التطورات التكنولوجية إمكان جمع

المعلومات وتجهيز البيانات بقدر أكبر من الكفاءة، وزيادة الإلمام بالحالة، وسرعة أكبر في اتخاذ القرار، والاشتباك بدقةٍ أكبر ولفترة أطول. وتتزايد أهمية المنظومات المسيَّرة الذاتية التشغيل التي تشغَّل من بُعد في الحروب الحديثة، وستغير هذه المنظومات مستقبل الحرب والعمليات وميدان المعركة. وستزداد أهمية توقُّع التطورات التكنولوجية، ودمج التكنولوجيات الناشئة في منظومات الدفاع والاستفادة مما هو غير متوقع، مع تسارع وتيرة التطور التكنولوجي في المستقبل. كما يمكن للتفوق التكنولوجي أن يعوض النقص العددي.

التحديات

في الوقت نفسه، من المهم تكوين فهم واسع للتهديدات الأمنية واحتمالات إساءة الاستخدام وقضايا حقوق الإنسان وأوجه الترابط المتعلقة بتطوير التكنولوجيات الثورية مثل الذكاء الاصطناعي. وستطرح هذه التكنولوجيات، في سياق تطورها، تحديات جديدة في قطاعي الدفاع والأمن على وجه الخصوص. إن تطور الذكاء الاصطناعي يجعل الهجمات السيبرانية وأنشطة التأثير على المعلومات، وأحد أدواتها، التضليل الإعلامي، أكثر استهدافا وفعالية. وعلاوة على ذلك، إن الذكاء الاصطناعي يُستخدم بالفعل للتأثير على الانتخابات. وفي هذه البيئة، يجب أيضا زيادة التركيز على حفظ أمن المعلومات السرية.

إن القانون الدولي، ولا سيما ميثاق الأمم المتحدة والقانون الدولي لحقوق الإنسان والقانون الدولي الإنساني، يسري بالكامل على الفضاء السيبراني. ويظل احترام إطار السلوك المسؤول للدول في الفضاء السيبراني والتقيد به أساسيين لصون السلام والأمن والاستقرار على الصعيد الدولي. فالتطور التكنولوجي يطرح قضايا جديدة متصلة مثلا بالبيئة السيبرانية، واستخدام الذكاء الاصطناعي، وتكنولوجيات الأسلحة الجديدة، واستغلال المواد الخام الحرجة. وقد تشمل أنشطةُ التأثير المختلطة ممارسات تهدف إلى إعاقة تحقيق المساءلة بموجب القانون الدولي. لذا، تنادي فنلندا بأخذ الحقوق الأساسية وحقوق الإنسان والمخاطر المرتبطة في الاعتبار بقوة عند تطوير وتطبيق الذكاء الاصطناعي ووضع القواعد الناظمة ذات الصلة. ومن المهم وضع مبادئ ومعايير وقواعد وسياسات وأطر وطنية تضمن التطبيقات المسؤولة للذكاء الاصطناعي في المجال العسكري، بما يتوافق مع القانون الدولي.

وقد أتاح التطور التكنولوجي فرصا جديدة للجهات الفاعلة المعادية كي تنخرط في أنشطة التأثير المختلطة التي تمارَس تحت عتبة النزاع المفتوح. وباتت العمليات السيبرانية العدائية جزءا راسخا من السياسة الخاضعة للقوة ومن مجموعة الأدوات المتاحة لأنشطة التأثير التي تقوم بها الجهات الحكومية. كما تُجرى العمليات السيبرانية والهجينة والمتعلقة بالمعلومات في ظروف عادية، وهي قد تحجب الحدود التي تفصل الحرب عن السلام. ورغم الطابع التكنولوجي المتزايد للحرب، فإن قدرات الحرب التقليدية لا تزال تتسم بالأهمية لا سيما في النزاعات الواسعة النطاق والطويلة الأمد.

وتواجه بلدان عدة أنشطة مكثفة للتأثير على المعلومات تستخدم الذكاء الاصطناعي هي أيضاً. وبات الاستخدام الضار للمعلومات جزءا يوميا من التأثير الواسع النطاق، وازدادت المنافسة في بيئة المعلومات.

ويوفر كلّ من التطورات في مجالي البنية التحتية والتكنولوجيا والعدد المتزايد من المستخدمين فرصا أكبر للقيام بأعمال عدائية في المجال السيبراني. ويواجه العديد من البلدان بشكل مستمر عمليات جمع المعلومات الاستخبارية على شبكات المعلومات، والتجسس السيبراني والهجمات السيبرانية من قبل جهات معادية تجهد للتأثير بشكل مادي على البنية التحتية الحيوية. وإلى جانب دور الجهات الحكومية،

يتنامى دور الجهات غير التابعة للدول ذات الدوافع السياسية أو التي تقودها الدول بوصفها جهات منظِّمة للأنشطة العدائية.

فرنسا

[الأصل: بالفرنسية]
[11 نيسان/أبريل 2025]

أولا - تأثير الذكاء الاصطناعي في المجال العسكري على السلم والأمن الدوليين

الفرص المتاحة

المساعدة في التخطيط واتخاذ القرارات - تعمل القوات المسلحة الفرنسية على استغلال قواعد بياناتها المتعلقة بالأحداث المرتبطة بالذخائر والمتفجرات من أجل تطوير أدوات للتنبؤ بالتهديدات المحتملة في منطقة محددة.

دعم الأشخاص - يُطرح نظام الذكاء الاصطناعي لتدريب أطقم الطيران لتحسين تدريب الطيارين الفرنسيين، من خلال تحليل البيانات التي يتم جمعها من الرحلات الجوية أو محاكاة الطيران. ويمكن للذكاء الاصطناعي أيضًا أن يساعد البشر في التعامل مع كميات كبيرة من البيانات، مثل نظام "الأذن الذهبية" الذي يعالج البيانات الصوتية بشكل كبير لتركيز انتباه المشغلين الفرنسيين على الإشارات ذات القيمة المضافة.

مواجهة نقاط ضعفنا في مجال تكنولوجيا المعلومات والاتصالات - يمكن استخدام تكنولوجيات الذكاء الاصطناعي في الأمن السيبراني ومواجهة انتشار المعلومات الخاطئة. وتعتمد القوات المسلحة الفرنسية على أنظمة الكشف عن التزييف العميق.

تعزيز تنفيذ القانون الدولي الإنساني وحماية الأشخاص والأعيان المشمولين بالحماية - يمكن أن يساهم الذكاء الاصطناعي في تنفيذ المبادئ الأساسية للقانون الدولي الإنساني، مثل التمييز والتناسب ومبدأ التحوط. كما يمكن استخدام الذكاء الاصطناعي لحماية الناس من خلال المساعدة في إزالة الألغام الأرضية باستخدام طائرات مسيرة مزودة بأجهزة استشعار تقترن بالذكاء الاصطناعي.

تعزيز تحديد الأسلحة - يمكن استخدام الذكاء الاصطناعي لرصد واكتشاف عمليات الإطلاق السرية أو التغييرات في مواقع إنتاج الأسلحة أو اختبارات الأسلحة الكيميائية والبيولوجية بشكل أفضل. وللتحكم في صادرات الأسلحة، يمكن للذكاء الاصطناعي تحسين إمكانية تتبعها.

تعزيز الوقاية وحفظ السلام وبناء السلام - من شأن الذكاء الاصطناعي أن يمكّن عمليات حفظ السلام من التكيف بشكل أفضل وأن يجعلها من ثم أكثر فعالية. وقد صُمم نظام "Resistance" للترجمة الفورية الذي طورته القوات المسلحة الفرنسية لتمكين التواصل مع السكان المحليين خارج شبكة الإنترنت وبدونها، وبالتالي مكافحة المعلومات المضللة.

المخاطر التي يجب التخفيف من حدتها

المخاطر الخاصة بالتكنولوجيا – تنطوي تقنيات التعلم على مخاطر مختلفة فيما يتعلق بالتحيز: التحيز غير المقصود؛ والتحيز التطوعي؛ والتحيز المتصل بإعادة تشكيل البيانات الحساسة بشكل خاص؛ والنتائج المبهمة أو التي يصعب تفسيرها. وهناك أيضًا مشكلة الاستهلاك المتسارع لموارد الطاقة.

تزايد المخاطر على الأمن والاستقرار الدوليين – في حال وقوع الذكاء الاصطناعي في أياد غير أمينة، يمكن أن يؤدي الذكاء الاصطناعي إلى تفاقم بعض المخاطر التي تهدد الأمن والاستقرار الدوليين (سيناريوهات التصعيد، وسباق التسلح، وانتشار الأسلحة باتجاه جهات فاعلة من غير الدول، وتوسيع نطاق عمليات النفوذ والأعمال العدائية في المجال السيبراني) والتي يجب التصدي لها بوضع التدابير الملائمة للتخفيف من المخاطر. وإن مخاطر عدم التمكين الناجمة عن الاعتماد على التكنولوجيا تستدعي ضمان مسؤولية البشر.

ثانيا – المبادئ والتدابير الرئيسية "للذكاء الاصطناعي المسؤول" طوال الدورة العمرية

تطوير ذكاء اصطناعي يحترم القانون الدولي الإنساني

تكييف استعراضات الشرعية – إذا كان هذا الاستعراض قابلا للتطبيق بشكل كامل على الذكاء الاصطناعي العسكري، فسيتعين تكييف طرائقه بدقة مع خصوصيات هذه التكنولوجيا.

إجراء الاستعراضات المناسبة – يجب إجراء هذا الاستعراض، حسب الضرورة، خلال مختلف مراحل الدورة العمرية لمنظومات الأسلحة. ويجب الاحتفاظ به عندما تخضع التدابير نفسها لابتكارات أو عندما تتضمن مكونات جديدة من المحتمل أن تعدل بشكل كبير من التأثيرات الناجمة عنها.

تطوير ذكاء اصطناعي موثوق وآمن

تقييم النظم وتأهيلها واعتمادها – يجب تقييم هذه النظم وتأهيلها لتصل إلى المستوى المناسب (وفقًا لأهمية الوظائف الحاسمة)، باستخدام تحليل المخاطر أثناء مرحلة التصميم. ويجب أن تكون مرتبطة بحالات استخدام محددة. ويجب مراجعة هذه الفحوصات على فترات زمنية تتناسب مع التحديات المطروحة.

الاعتماد على البيانات السيادية الخاضعة للرقابة – يجب وضع تدابير مضادة للتعامل مع مخاطر اختراق البيانات، ويجب وضع الدفاعات المناسبة.

تصحيح النظم وإعادة تدريبها – من المهم تحديد وتوصيف أي أخطاء تتم مواجهتها (أثناء الاختبار أو الاستخدام التشغيلي)، لتوعية المشغلين بالحاجة إلى الدروس المستفادة والتحقق باستمرار من أن النظام يتوافق مع التزاماتنا الدولية.

إخضاع الذكاء الاصطناعي لتحكم بشري مناسب وتسلسل قيادي مسؤول

ضمان توافق القرارات والإجراءات مع القانون – يجب أن يكون المشغل أو القائد العسكري قادرًا على ممارسة سلطته التقديرية للتحقق من مدى توافق النتائج المقترحة مع الأوامر المعطاة ومع الالتزامات القانونية.

ضمان مسؤولية البشر – إن المسؤولية البشرية في تصميم تكنولوجيات الذكاء الاصطناعي ونشرها واستخدامها مبدأ غير قابل للتصرف، مما يتطلب إضفاء الطابع الرسمي على تسلسل المسؤولية لدى المسؤولين عن وظائف القيادة والتحكم والتنفيذ.

تكييف التحكم البشري – يعتبر تحليل وتوصيف التحكم البشري المناسب، دون تقييد قدرات النظام الذي يدمج تكنولوجيات الذكاء الاصطناعي مسألة معقدة يجب أن تأخذ في الاعتبار مختلف العوامل البشرية والتقنية والسياقية.

تدريب القادة العسكريين والأفراد العسكريين على إتقان هذه النظم – يجب إدخال مرحلة تدريب وممارسة قبل الاستخدام، لتوعية الموظفين بالفوائد والمخاطر.

تطوير ذكاء اصطناعي مستدام

حماية الأبحاث – يجب أن تكون البرامج البحثية مفتوحة في نطاقها وأفقها، دون فرض محظورات فضفاضة بشكل مفرط مسبقاً.

إخضاع البحث للتفكير الأخلاقي – أنشأت فرنسا هيئة دائمة، وهي لجنة أخلاقيات الدفاع، لدراسة القضايا الأخلاقية المتعلقة بالتكنولوجيات الجديدة في قطاع الدفاع.

تطوير ذكاء اصطناعي مقتصد – إن تفضيل السلوك المقتصد يعني التفكير في استخدام الذكاء الاصطناعي وتحسين مرونة واستدامة هذه النظم، مع التحكم في التكاليف.

ثالثا – عملية مخصصة لوضع حوكمة عالمية تهدف إلى تفعيل مبادئ الذكاء الاصطناعي المسؤول

عملية عالمية وشاملة – يجب أن تضم المناقشات جميع أصحاب المصلحة: الدول – تجدر الإشارة على وجه الخصوص إلى أن المشاركة الفعالة للدول التي تطور وتستخدم هذه النظم أمر ضروري للغاية؛ وفي ضوء هذه الضرورة الحتمية، سيتعين على عملية اتخاذ القرار أن تأخذ بعين الاعتبار مختلف المواقف وتعتمد قواعد لهذا الغرض بهدف ضمان توافق الآراء – وكذلك قطاع الصناعة والأوساط العلمية والأكاديمية والمجتمع المدني، وذلك لتجنب المناقشات المنفصلة عن الواقع والحفاظ على الابتكار. ويمكن أن تشكل اللجنة الأولى للجمعية العامة منتدى مناسباً.

هيكل حوكمة مبسط ومتماسك – ينبغي أن يتيح إطار عمل واحد ترشيد الجهود المبذولة لتحسين الكفاءة وتعزيز فرص تحقيق الأثر المنشود. وسيكون من الضروري ضمان التكامل مع مناقشات فريق الخبراء الحكوميين المعني بالتكنولوجيات الناشئة في مجال منظومات الأسلحة الفتاكة الذاتية التشغيل، الذي يجب أن يكون قادراً على مواصلة عمله بعد عام 2026، بموجب ولاية جديدة.

عملية ذات منحى عملي، تركز على القضايا الخاصة بالمجال العسكري – يجب أن تركز الحوكمة على القوانين المنطبقة على النزاعات المسلحة، وبالتالي يجب أن تركز أولاً وقبل كل شيء على القانون الدولي الإنساني. ويجب أن تكون أولوية أي عملية دولية هي ضمان الامتثال للقواعد القانونية القائمة، من خلال مناقشة وضع مبادئ توجيهية وسبل تنفيذها من قِبل الدول (عن طريق تيسير تبادل أفضل الممارسات وتعزيز التعاون والمساعدة الدوليين، بطرق تتناسب مع الشؤون العسكرية)، مع تعزيز التدابير المناسبة لبناء الثقة والحد من المخاطر.

ألمانيا

[الأصل: بالإنكليزية]
[11 نيسان/أبريل 2025]

أولا - مقدمة

شهدت السنوات الأخيرة تطوراً غير مسبوق في تكنولوجيا الذكاء الاصطناعي بما في ذلك تطوير التطبيقات القائمة على التكنولوجيات الثورية مثل الذكاء الاصطناعي التوليدي. ولا غنى للدول عن الاستفادة من الفرص الناشئة عن هذه التطورات التكنولوجية وضمان عدم إعاقة التقدم التكنولوجي. بيد أنه في الوقت نفسه يتعين على الدول أن تضمن تطوير واستخدام تطبيقات الذكاء الاصطناعي في المجال العسكري بشكل مسؤول ومتوافق تماماً مع القانون الدولي، بما في ذلك القانون الدولي الإنساني. وفي هذا الصدد، يتسم التبادل الدولي بأهمية قصوى من أجل إتقان تحقيق هذا التوازن.

وإزاء هذه الخلفية، تنشط ألمانيا في الإسهام في العمليات الدولية بشأن المسائل المتصلة بالاستخدام المسؤول للذكاء الاصطناعي في المجال العسكري. وقد روجت ألمانيا لجملة أمور منها قرار الجمعية العامة 239/79 بشأن الذكاء الاصطناعي في المجال العسكري وآثاره على السلام والأمن الدوليين بصفتها عضوا في الفريق الأساسي لمقدمي مشروع القرار، وهي تؤيد بالكامل جهود الأمين العام الرامية إلى تقديم تقرير موضوعي عن آراء الدول الأعضاء بشأن "الفرص والتحديات التي يطرحها تطبيق الذكاء الاصطناعي في المجال العسكري على السلام والأمن الدوليين".

وترحب ألمانيا بفرصة الوقوف على آراء الدول الأعضاء والجهات المعنية الأخرى بمزيد من التعمق وبطرح عناصر اعتباراتها الخاصة في سياق تناوُل هذه المسائل الهامة.

ثانيا - المبادئ وافتراضات العمل

يعتمد نهج ألمانيا في ضمان الاستخدام العسكري المسؤول للذكاء الاصطناعي على المبادئ الأساسية التالية التي حُددت في إطار المنتديات والمناقشات الدولية المختلفة.

وقد أسهمت ألمانيا بفعالية في وضع مبادئ منظمة حلف شمال الأطلسي (الناتو) للاستخدام المسؤول في عام 2021، وهي لا تزال متقيدة تماما بتلك المعايير الهامة: الشرعية في تطوير واستخدام تطبيقات الذكاء الاصطناعي؛ المسؤولية البشرية التي تضمن المساءلة في تصميم وتشغيل الذكاء الاصطناعي في المنظومات العسكرية؛ قابلية شرح وتتبّع تطبيقات الذكاء الاصطناعي في المجال العسكري؛ ضمان الموثوقية والسلامة والأمن والطابع المعزَّز طوال دورة عمر المنظومات التي تعمل بالذكاء الاصطناعي والمزودة بوظائف التشغيل الذاتي؛ إمكانية الحوكمة، عبر ضمان إتاحة التفاعل المناسب بين الإنسان والآلة والتقليل من التحيز.

وإضافة إلى ذلك، أيدت ألمانيا الوثائق الختامية لقمتيّ تسخير الذكاء الاصطناعي المسؤول في المجال العسكري اللتين عقدتا في لاهاي عام 2023 (دعوة إلى العمل) وسُول في عام 2024 (مخطط للعمل)، وكذلك الإعلان السياسي بشأن الاستخدام العسكري المسؤول للذكاء الاصطناعي والتشغيل الذاتي الذي بادرت به الولايات المتحدة الأمريكية في عام 2023، وهي تنشط في تنفيذ الإعلان.

وعلاوة على ذلك، إن ألمانيا هي أيضا جزء من مبادرة "شراكة الذكاء الاصطناعي من أجل الدفاع"، التي تشجّع فيها الدولُ المتماثلةُ التفكير على الاستخدام المسؤول للذكاء الاصطناعي، وتعزيز المصالح المشتركة وأفضل الممارسات في تطبيق الأخلاقيات المتعلقة بالذكاء الاصطناعي، ووضع أطر لتيسير التعاون وتنسيق توجيه الرسائل الاستراتيجية بشأن سياسة الذكاء الاصطناعي.

وفي شباط/فبراير 2025، صادقت ألمانيا على إعلان باريس بشأن الإبقاء على التحكم البشري في منظومات الأسلحة المدعومة بالذكاء الاصطناعي، مسلطةً بذلك الضوءَ على أهمية صون التحكم البشري في تطبيق الذكاء الاصطناعي في المجال العسكري.

ثالثا - الجوانب الرئيسية المتعلقة باستخدام الذكاء الاصطناعي في القوات المسلحة الاتحادية الألمانية

تدرس القوات المسلحة الاتحادية الألمانية إمكان استخدام الذكاء الاصطناعي لإنجاز مهمتها الأساسية وتحقيق التفوق من حيث المعلومات واتخاذ القرار والفعالية، وكذلك لتنفيذ الآليات العمليات الإدارية واللوجستية وتلك المتعلقة بالصيانة التنبؤية للمنظومات المعقدة على النحو الأمثل. ويُستخدم الذكاء الاصطناعي أيضا لدعم الموظفين المتخصصين في سياق الكشف المبكر عن الأزمات المدنية والعسكرية في مجالات مختلفة في تحليل البيانات الجماعية وإجراء التوقعات لعمليات الانتشار. والذكاء الاصطناعي جزء لا يتجزأ من مشاريع الدفاع الكبرى التي تنفَّذ أيضا في سياق أوروبي، ما يسهم في الإبقاء على التميز التكنولوجي الأوروبي وتعزيزه. وفي ما يتعلق بالتطورات الوطنية والتكنولوجية في قطاع التسلح على النطاق الدولي، يُستخدم الذكاء الاصطناعي لضمان القدرات المطلوبة للدفاع الوطني والحليف في المستقبل. ويجرى العمل على تطوير احتمالات نشر الذكاء الاصطناعي، لا سيما لحماية الأمن القومي وللأغراض العسكرية، ضمن اختصاصات ومسؤوليات الوزارات والإدارات المعنية. ومن دون الإخلال بما سبق، أُدرجت تكنولوجيا الذكاء الاصطناعي وتطبيقات الذكاء الاصطناعي ذات الصلة بالأمن ضمن استراتيجية الذكاء الاصطناعي للحكومة الاتحادية الألمانية.

وتشترط القوات المسلحة الاتحادية الألمانية أعلى المطالب الأخلاقية وتضع أعلى المعايير القانونية لاستخدام الذكاء الاصطناعي في منظومات الأسلحة. وعلى وجه الخصوص، تتبع القوات المسلحة الاتحادية الألمانية أحكامَ القانون الدولي الإنساني في ما يتعلق بالنزاعات المسلحة والمبادئَ التوجيهية للجنة أخلاقيات البيانات التابعة للحكومة الاتحادية والناتو ولا سيما المبادئ الست المذكورة أعلاه للاستخدام المسؤول للاستخدام العسكري للذكاء الاصطناعي، طوال دورة حياة تلك المنظومات.

رابعا - اعتبارات أساسية

بغية الحفاظ على القدرات الدفاعية والرادعة اللازمة، تظل ألمانيا مصممة على اغتنام الفرص المتعلقة بالذكاء الاصطناعي في المجال العسكري، وهي مقتنعة بوجوب عدم إعاقة التقدم التكنولوجي، لا سيما بالنظر إلى الطابع المزدوج الاستخدام الملازم للتكنولوجيات المعنية.

وفي الوقت نفسه، ستواصل ألمانيا توسيع قاعدة المعارف من خلال تقييم ومواجهة المخاطر المرتبطة باستخدام الذكاء الاصطناعي في المجال العسكري، بما في ذلك تلك المتصلة بأوجه التحيز غير المقصودة، كالتحيز الجنساني. وفي هذا السياق، تعلق ألمانيا أهمية كبيرة على الدور الأساسي للأوساط الأكاديمية وللإسهامات القيّمة التي تقدمها معاهد البحوث ومراكز الفكر العاملة في هذا المجال. وبغية

تشجيع البحوث ذات الصلة، تدعم ألمانيا المنظمات البحثية المختصة ومن ضمنها معهد الأمم المتحدة لبحوث نزع السلاح، عبر تقديم مساهمة مالية في مشاريع بحثية محددة الأهداف.

ومن الأهمية بمكان بالنسبة إلى ألمانيا ضمان أن تكون المناقشات شاملة للجميع، سواء من الناحية الجغرافية أو عبر مراعاة آراء الدول الأعضاء، فضلا عن آراء القطاع المعني والمجتمع المدني والأوساط الأكاديمية.

كما تولي ألمانيا عند تناولها الفرص والمخاطر المرتبطة بمنظومات الأسلحة المدعومة بالذكاء الاصطناعي أهميةً خاصة لمفهوم التحكم البشري وتعتبر وجود إطار فعال للتحكم البشري شرطا أساسيا لضمان امتثال كل منظومات الأسلحة للقانون الدولي الإنساني. ولا يشمل ذلك التحكمَ التقني فحسب بل أيضا عنصر التقدير البشري. ويشمل مفهوم ألمانيا لإطار التحكم البشري مجموعة من الخطوات والإجراءات الممكنة من الناحية التكنولوجية التي ترسم حدودا واضحة يُسمح لخوارزمية المنظومة بالعمل ضمنها. ويشكل القانون الدولي، ولا سيما القانون الدولي الإنساني، عنصراً أساسياً ضمن هذه الحدود. وعندما يتعلق الأمر بالاستخدام الفعلي للذكاء الاصطناعي في ساحة المعركة، يتسم السياق بأهمية قصوى. وترى ألمانيا أن استحداث مفهوم إطار التحكم البشري وسيلة مناسبة لمراعاة هذا الأمر بشكل كافٍ.

ومن الضروري إيلاء اهتمام خاص عندما يكون استخدام الذكاء الاصطناعي متصلا بالأسلحة النووية، وهو مجال لا يزال فيه النقاش العلمي والسياسي في مراحله الأولى. وقد تكون للاستخدام المحتمل للذكاء الاصطناعي في أنظمة القيادة والتحكم في الأسلحة النووية تداعيات خطيرة على الاستقرار الاستراتيجي أو التصعيد النووي. وفي الوقت نفسه، قد يفتح الذكاء الاصطناعي سبلًا جديدة لاحتواء انتشار واستخدام أسلحة الدمار الشامل. وقد سعت ألمانيا إلى الإسهام في هذه المناقشات باستضافة مؤتمر حول الذكاء الاصطناعي وأسلحة الدمار الشامل وذلك في إطار سلسلة مؤتمراتها الثابتة حول موضوع "السيطرة على التكنولوجيا – إعادة التفكير في تحديد الأسلحة"، عُقد في برلين في 28 حزيران/يونيه 2024.

إن اتفاقية حظر استحداث وإنتاج وتخزين الأسلحة البكتريولوجية (البيولوجية) والتكسينية وتدمير تلك الأسلحة واتفاقية حظر استحداث وإنتاج وتخزين واستعمال الأسلحة الكيميائية وتدمير تلك الأسلحة تحظران فئات كاملة من أسلحة الدمار الشامل. ويمكن لتطبيقات مثل النماذج اللغوية الكبيرة (التوليدية) أن تسهّل انتشار المعارف ذات الاستخدام المزدوج التي قد يُساء استخدامها لتطوير أو إنتاج أو استخدام الأسلحة البيولوجية والكيميائية. كما يمكن أن يؤدي التقارب بين تطبيقات الذكاء الاصطناعي مثل AlphaFold والبيولوجيا التركيبية إلى تمكين جهات خبيثة من تصميم بروتينات جديدة يمكن أن تفلت من الكشف بسبب التغييرات في تسلسل الحمض النووي. ويمكن استخدام الذكاء الاصطناعي لتحليل سحُب البيانات الضخمة، مثل بيانات الجينوم البشري، كما يمكنه أن يحقق فوائد كبيرة لتطوير علاجات طبية فردية، ولكن يمكن في المقابل أن يُساء استخدامه لتطوير أسلحة بيولوجية تستهدف مجموعات عرقية محددة.

لذا، ستواصل ألمانيا، بتعاون وثيق مع شركائنا الدوليين، تحديد خطوط العمل الممكنة لتقييم تأثير تطبيقات الذكاء الاصطناعي على تطوير وإنتاج الأسلحة المحظورة واعتماد القواعد الناظمة الممكنة. وفي الوقت نفسه، ستستفيد ألمانيا من مزايا الذكاء الاصطناعي في مجالات التحقق والتحليل الجنائي البيولوجي والحد من المخاطر.

خامسا – التزام ألمانيا بالآليات الدولية

ما برحت ألمانيا تسهم بنشاط في آلية الذكاء الاصطناعي المسؤول في المجال العسكري منذ إنشائها، وهي ستواصل القيام بذلك. وألمانيا هي من ضمن مجموعة الدول المشاركة في تقديم قرار الجمعية العامة 239/79 بشأن الذكاء الاصطناعي في المجال العسكري وآثاره على السلام والأمن الدوليين. وتنوّه ألمانيا بشدة بالنهج الأقاليمي والمتعدد الجهات المعنية المنبثق من هذه المبادرة الهامة وتتطلع إلى مواصلته في إسبانيا في أيلول/سبتمبر 2025.

وفي إطارٍ من التكامل التام، أسهمت ألمانيا في الإعلان السياسي بشأن الاستخدام العسكري المسؤول للذكاء الاصطناعي والمنظومات الذاتية التشغيل الذي بادرت به الولايات المتحدة بما في ذلك عبر المشاركة (مع النمسا) في رئاسة الفريق العامل المعني بالرقابة.

وعلاوة على ذلك، فإن ألمانيا منخرطة بفعالية في شراكة الذكاء الاصطناعي من أجل الدفاع وهي تشارك في شبكة خبراء معهد الأمم المتحدة لبحوث نزع السلاح بشأن حوكمة الذكاء الاصطناعي في المجال العسكري.

كما تدعم ألمانيا رئيسَ فريق الخبراء الحكوميين المعني بالتكنولوجيات الناشئة في مجال منظومات الأسلحة الفتاكة الذاتية التشغيل في جنيف السفير روبرت إن دن بوش، وتواصل المشاركة بنشاط في هذه العملية، بما في ذلك عبر تنسيق مواقف العديد من الدول الأعضاء في إطار ما يسمى الفريق ذي المستويين. وبتعاون وثيق مع شركائنا الدوليين، ستواصل ألمانيا العمل من أجل الوفاء بولاية الفريق في الوقت المناسب، ويفضَّل أن يكون ذلك بحلول نهاية عام 2025.

وفي سياق الناتو، تعي ألمانيا إمكانات الذكاء الاصطناعي في مواصلة تطوير القوات المسلحة والقدرات الدفاعية للحلف، فضلاً عن التحديات التي سيطرحها استخدام الذكاء الاصطناعي بالنسبة إلى قابلية التشغيل البيني للقوات المسلحة للدول التي يتكوّن منها الحلف. ويجب أن تؤخذ تطورات الذكاء الاصطناعي المتعددة الجنسيات وجوانب توحيد معايير الذكاء الاصطناعي داخل الناتو والاتحاد الأوروبي والبلدان الشريكة لألمانيا في الاعتبار بشكل كامل، من أجل ضمان قابلية التشغيل البيني للقوات المسلحة الاتحادية الألمانية كقوة عسكرية في سياق العمليات الدولية. ونتيجة لذلك، رحبت ألمانيا باتفاق بلدان الناتو على مبادئ الاستخدام المسؤول في سياق استراتيجية الناتو بشأن الذكاء الاصطناعي.

سادسا – سبل المضي قدما

مع استمرار تطور التكنولوجيات الثورية الناشئة وتأثيرها على عالمنا، تعتبر ألمانيا أن التنسيق الدولي الشامل للجميع في التطوير العسكري المسؤول للذكاء الاصطناعي واستخدامه أمر لا غنى عنه. وتوفر الآليات الدولية الحالية إطاراً ممتازاً لمعالجة الجوانب الهدفة التي ينطوي عليها الأمر ولمراعاة آراء مجموعة متنوعة من الجهات المعنية. وستواصل ألمانيا الإسهام بفعالية في هذه الجهود من أجل تنفيذ وتوسيع نطاق الدعم للالتزامات السياسية بشأن الاستخدام العسكري المسؤول للذكاء الاصطناعي مثل الإعلان السياسي الذي تقوده الولايات المتحدة وآلية الذكاء الاصطناعي المسؤول في المجال العسكري. وتتطلع ألمانيا إلى درس نتائج تقرير الأمين العام عن الذكاء الاصطناعي في المجال العسكري. وستواصل

ألمانيا الإسهام بفعالية في العملية المتعلقة بمنظومات الأسلحة الفتاكة الذاتية التشغيل في إطار فريق الخبراء الحكوميين في جنيف.

اليونان

[الأصل: بالإنكليزية]
[10 نيسان/أبريل 2025]

أثّر دمج الذكاء الاصطناعي في قطاع الدفاع بشكل أساسي على وسائل وأساليب تنفيذ العمليات العسكرية. فالتطبيقات العسكرية للذكاء الاصطناعي تقدم فوائد عملياتية جمّة بينها تسريع اتخاذ القرار، وتعزيز كشف التهديدات والتنبؤ بها، والإلمام بالأوضاع وتقييمها بشكلٍ آني، وتخصيص الموارد والتخطيط على النحو الأمثل، وتقديم الدعم اللوجستي، وتعزيز القدرات البشرية في المهام المعقدة، والتجهيز الفعال للبيانات الاستخبارية الواسعة النطاق.

بيد أنه رغم هذه التطورات، من الضروري الإقرار بأن التقدم التكنولوجي يحمل معه أيضا تحديات معقدة ومتعددة الأبعاد تتطلب تمحيصا متأنيا يضمن عدم تقويض السلام والأمن والاستقرار على الصعيدين الإقليمي والعالمي.

وفي هذا الصدد، وبالنسبة إلى اليونان، يشكل استخدام المنظومات العسكرية المزودة بقدرات التعلم الآلي أحد المجالات الرئيسية المثيرة للقلق إذ إنه يطرح عددا من التحديات منها الشفافية وقابلية التفسير، لأن النماذج المعقدة يمكن أن تعمل كـ "صناديق سوداء" في سياق آليات غير محددة لاتخاذ القرار، لا سيما في ظل بيئة متطورة باستمرار لساحة المعركة.

وإضافة إلى ذلك، فإن الاستخدام المحتمل للذكاء الاصطناعي التوليدي في العتاد العسكري يضفي طبقة هامة من التعقيد وعدم اليقين، إذ قد تولّد هذه المنظومات من تلقاء نفسها حلولاً جديدة وتتكيف مع ظروف ساحة المعركة المتغيرة من خلال تحليل البيانات الجديدة والتعلم منها بشكل مستمر – وهي قدرات تثير قلقا بالغا لليونان. ولمواجهة هذه التحديات، من الأهمية بمكان فرض حدود وقيود عملياتية واضحة على استخدام هذه المنظومات للحؤول دون سلوك غير مقصود.

وبالنظر إلى السياق المذكور أعلاه، فإن واحدا من أكثر التحديات المثيرة للقلق في ما يتعلق باستخدام الذكاء الاصطناعي في السياقات العسكرية يكمن في إدماجه في منظومات القيادة والتحكم ودعم اتخاذ القرار في استخدام الأسلحة النووية. إن احتمال تفويض القرارات المتعلقة بالردع النووي، أو حتى مباشرة تطبيق البروتوكولات ذات الصلة باستخدام تلك الأسلحة، لمنظومات مدعومة بالذكاء الاصطناعي يستدعي دراسة متأنية تضمن الإشراف والانخراط البشريين لدى اتخاذ تلك القرارات ووضع ضمانات مناسبة في مجال الأمن السيبراني تمنع التصعيد غير المقصود.

وما يثير القلق بالقدر عينه، في ظل البيئة الجيوسياسية الحالية الملأى بالتحديات هو الجهد الذي تبذله الدول للحفاظ على تفوقها العسكري، وهو جهد يمكن أن يؤجج سباق تسلحٍ يتسم بانعدام الشفافية والشك المتبادل. ويمكن أن تؤدي هذه المنافسة إلى تفاقم حال عدم الاستقرار الجيوسياسي وإلى طرح تحديات كبيرة على الأمن العالمي في ظل اختلال توازن القوى وازدياد الاتساع الواضح للفجوة التكنولوجية بين الدول المتقدمة والنامية.

وعلاوة على ذلك، فإن التطوير والنشر المتزايدين للقدرات المدعومة بالذكاء الاصطناعي من قبل القوات المسلحة من شأنهما خفض عتبة النزاع المسلح. فالوتيرة المتسارعة في اتخاذ القرارات والاعتماد المتزايد على المنظومات المسيَّرة في مسارح العمليات يزيدان من مخاطر التصعيد غير المقصود، حيث يُستبدل بشكل متزايد العنصر البشري في ساحة المعركة بمنظومات مسيَّرة.

وفي هذا السياق، ثمة معيار آخر يستدعي النظر فيه على النحو المناسب هو انتشار القدرات المدعومة بالذكاء الاصطناعي وتحويلها لصالح دول تتجاهل النظام الدولي القائم على القواعد ولجهات من غير الدول، بينها المنظمات الإرهابية. وبازدياد إمكانية الوصول إلى تكنولوجيا الذكاء الاصطناعي، يبرز خطر كبير في أن تتمكن هذه الجهات من الحصول عليها ونشرها لتحقيق أهداف مزعزعة للاستقرار، الأمر الذي يزيد من التحدي المطروح على الأمن الدولي.

كما تخلق التطبيقات العسكرية للذكاء الاصطناعي مخاطر وتحديات تتعلق بالعمليات النفسية والمعلومات المضللة، إذ إنها تتيح إنتاج معلومات مغلوطة وعمليات تزييف بالغ وبيانات مزورة على نطاق واسع تهدف إلى خداع الجمهور وزعزعة استقرار المؤسسات. وتعمل الحسابات المؤتمتة (البوتات) وخوارزميات الدعاية الموجَّهة على تدعيم العمليات النفسية بتأثيرها على الرأي العام والعمليات الانتخابية وخلق التوترات الاجتماعية، بما في ذلك تقويض ثقة السكان في عمليات حفظ السلام من خلال حملات التضليل الإعلامي. كما أن التحيزات الاجتماعية، كتلك المتعلقة بالجنس والعمر والعرق والإعاقة، تثير مخاوف هي أيضا، ومن الضروري إجراء تقييمات للمخاطر وتنفيذ تدابير للتخفيف من المخاطر منعاً للتحيز غير المقصود والتمييز في الخوارزميات.

وعلاوة على ذلك، يمكن استخدام تطبيقات الذكاء الاصطناعي في مجال الأمن السيبراني لحماية البنى التحتية الحيوية أو لأغراض خبيثة، مثل شن الهجمات السيبرانية واعتراض البيانات. وتتطلب التهديدات الهجينة التي تجمع بين العمليات العسكرية التقليدية والأساليب الاستخبارية الهجومية مزيداً من اليقظة والتنسيق بين الدول والجهات الفاعلة الدولية لتجنب التصعيد والحفاظ على السلام والأمن الإقليميين والدوليين.

وفي ضوء ما سبق، تؤيد اليونان بقوة الجهود الدولية الرامية إلى ضمان الاستخدام المسؤول للذكاء الاصطناعي في المجال العسكري، فهو، رغم التحديات المذكورة أعلاه، يمكن أن يعزز تطبيقَ القانون الدولي الإنساني ويسهم في حماية المدنيين بتحسين دقة إصابة الأهداف وتعزيز المراقبة والتحسين الأمثل لتقديم المساعدة الإنسانية.

ومن هذا المنطلق، نظمت اليونان في 4 نيسان/أبريل 2025، بالاشتراك مع جمهورية كوريا وفرنسا، وبدعم قيّم من أرمينيا وإيطاليا ومملكة هولندا، اجتماعاً لمجلس الأمن بصيغة آريا حول موضوع "تسخير الذكاء الاصطناعي الآمن والشامل للجميع والجدير بالثقة من أجل صون السلام والأمن الدوليين". وقدم هذا الاجتماع رؤى متبصرة قيّمة حول السبل التي يمكن للأمم المتحدة أن تسهم من خلالها في صون السلام والأمن الدوليين، لا سيما عبر تنظيم قدرات الذكاء الاصطناعي في المجال العسكري وعدم انتشارها ومنع تحويلها لصالح جهات أخرى، وتعزيز سيادة القانون والقيم الديمقراطية والتماسك الاجتماعي والتنمية الاقتصادية.

وإضـــافة إلى ذلك، أيدت اليونان، في إطار انخراطها على المســـتوى الدولي، البيانين المشـــتركين الصـــادرين عن القمتين حول الذكاء الاصـــطناعي المسـؤول في المجال العسـكري، اللتين عُقدتا في لاهاي (15 و 16 شـــباط/فبراير 2023) وسـول (9 و 10 أيلول/سـبتمبر 2024) بشـأن إجراءات تطوير الذكاء الاصـطناعي واستخدامه المسؤول في المجال العسكري. وأيدت اليونان أيضـا الإعلان السياسـي الذي تقوده الولايات المتحدة الأمريكية بشأن الاستخدام المسؤول للذكاء الاصطناعي والمنظومات الذاتية التشغيل وإعلان باريس بشأن الإبقاء على التحكم البشري في منظومات الأسلحة المدعومة بالذكاء الاصطناعي.

وعلاوة على ذلك، أنشـأت اليونان لجنة اسـتشـارية رفيعة المسـتوى[1] معنية بالذكاء الاصـطناعي لوضـع اسـتراتيجية وطنية شـاملة للذكاء الاصـطناعي، إلى جانب الهياكل اللازمة داخل وزارة الدفاع الوطني لمواجهة التحديات التكنولوجية والقانونية والأخلاقية والسـياسـية الناشـئة عن تطبيقات الذكاء الاصـــطناعي والمنظومات الذاتية التشغيل في المجال العسكري.

وأخيراً وليس آخراً، وبغية الإسهام على نحو بنّاء في الحوار الدولي حول الاستخدام المسؤول للذكاء الاصـطناعي في المجال العسكري، تعمل اليونان على تنظيم مؤتمر دولي حول موضـوع "النزاعات المسـلحة وإدارة الأزمات في عصر الذكاء الاصطناعي"، يُعقَد في أثينا يومي 22 و 23 أيار/مايو 2025.

الهند

[الأصل: بالإنكليزية]
[1 نيسان/أبريل 2025]

الذكاء الاصطناعي تكنولوجيا تحويلية تؤثر بشكل كبير على كل جانب من جوانب الحياة البشرية. وهو يشهد تطويرا بسرعةٍ وعلى نطاقٍ غير مسبوقين كما أنه يُعتمد ويُنشر بسرعة في طائفة من التطبيقات. ويمكن أن تكون للذكاء الاصـطناعي آثار تحويلية في الحد من الفقر وتحسين حياة الناس. ويتسم هذا الأمر بأهمية خاصة للبلدان النامية كالهند.

ومن الضـروري بذل جهود عالمية جماعية لوضـع حوكمة ومعايير للذكاء الاصـطناعي تدعم قيمنا المشـتركة وتتصـدى للمخاطر وتبني الثقة. وينبغي لحوكمة الذكاء الاصـطناعي ومعاييره: أن تراعي الترابط العميق عبر الحدود؛ وأن تشـجع الابتكار؛ وأن تُنشَـر بما يحقق مصـلحة العالم؛ وأن تعزز إمكانية الوصـول والإنصـاف بما يضـمن إتاحة فوائد الذكاء الاصـطناعي للجميع، لا سيما بلدان الجنوب. والهند ملتزمة بإجراء مناقشات مفتوحة حول الابتكار والحوكمة.

ويتعين إرساء المناقشات بشأن الذكاء الاصطناعي العسكري على الواقع العسكري، حيث يحصل دمج سريع للذكاء الاصـطناعي في العقائد والعمليات العسكرية. وقد أظهرت النزاعات الدائرة في أنحاء العالم المخاطر والفرص الناجمة عن الاعتماد المتزايد على هذه التكنولوجيات.

(1) تقدم الدراسة البارزة التي أجرتها اللجنة بعنوان "مخطط لتحوّل اليونان في مجال الذكاء الاصطناعي" مبادئ توجيهية ومشاريع رائدة لدفع عجلة التقدم في مجال الذكاء الاصطناعي في اليونان، وهي تتضمن أولويات تشمل صون الديمقراطية وتعزيزها، والتخفيف من آثار تغير المناخ والتكيف معه، ودعم الأمن.

ويطرح تطوير ونشر واستخدام الذكاء الاصطناعي في المجال العسكري تحديات أخلاقية وقانونية وأمنية. إن الهند، ومن دون التقليل من شأن هذه التحديات، تؤيد الرأي الذي يؤمن بقدرة الذكاء الاصطناعي على تحسين الامتثال للقانون الدولي الإنساني.

والهند تدعم الجهود العالمية الجماعية الرامية إلى تنظيم تطوير ونشر واستخدام الذكاء الاصطناعي في المجال العسكري بشكل مناسب. وينبغي لهذه الجهود أن تتناول الهواجس القانونية والأخلاقية وأن تتيح تحديدَ المخاطر المرتبطة بالذكاء الاصطناعي في المجال العسكري والتخفيفَ منها.

وينبغي لأي جهود جماعية في سبيل تنظيم الذكاء الاصطناعي في المجال العسكري بشكل مناسب أن تركز على التطبيقات والاستخدام، لا على التكنولوجيا ومكوناتها. وينبغي تجنب وصم تلك التكنولوجيا. كما يجب عدم تقييد الوصول إلى التكنولوجيا المعدّة لأغراض التنمية.

وينبغي استخدام الذكاء الاصطناعي بشكل قانوني في المجال العسكري، وفقًا للحق الراسخ في الدفاع عن النفس فرديا أو جماعيا بموجب القانون الدولي. ولا يزال القانون الدولي الإنساني يسري بالكامل على الذكاء الاصطناعي في المجال العسكري. والمبادئ الأساسية الواردة في القانون الدولي الإنساني، أي التمييز والتناسب والحيطة تنطبق على كل وسائل وأساليب الحرب في الماضي والحاضر والمستقبل.

إن التقدير والرقابة البشريين في استخدام الذكاء الاصطناعي في المجال العسكري ضروريان للتخفيف من المخاطر وضمان الامتثال للقانون الدولي الإنساني.

وينبغي لأي جهد جماعي أو قواعد ناظمة مناسبة في ما يتعلق بالذكاء الاصطناعي في المجال العسكري أن تراعي الالتزامات القانونية القائمة وأن تحترم الولاية القضائية والاختصاص الوطنيين، وكذلك القدرات الوطنية المعنية.

إن الهند ملتزمة بالاستخدام المسؤول للذكاء الاصطناعي في المجال العسكري.

كما أن الهند تعكف على وضع إطار لتقييم ذكاء اصطناعي جدير بالثقة في قطاع الدفاع لمواجهة التحديات المعقدة التي تطرحها تكنولوجيات الذكاء الاصطناعي الحديثة. ويتمحور هذا الإطار حول خمسة مبادئ رئيسية هي: (أ) الموثوقية والطابع المعزَّز؛ (ب) والسلامة والأمن؛ (ج) والشفافية؛ (د) والإنصاف؛ (هـ) والخصوصية. وتوفر هذه المبادئ أساسا لمزيد من المناقشات من أجل التنظيم المناسب لتطوير الذكاء الاصطناعي ونشره واستخدامه في المجال العسكري.

إندونيسيا

[الأصل: بالإنكليزية]
[11 نيسان/أبريل 2025]

ترحب إندونيسيا بمناقشة الفرص والتحديات التي يطرحها تطبيق الذكاء الاصطناعي في المجال العسكري على السلام والأمن الدوليين، مع التركيز بشكل خاص على مجالات أخرى غير الأسلحة الفتاكة الذاتية التشغيل، عملاً بالفقرتين 7 و 8 من قرار الجمعية العامة 239/79.

وبما أن الذكاء الاصطناعي في المجال العسكري يشمل طائفة واسعة من المنظومات والتطبيقات تستدعي إجراء مداولات متعددة الأطراف شاملة للجميع حول هذا الموضوع في الأمم المتحدة، فإنه ينبغي

لهذه المناقشة أن تطال ما هو أبعد من القدرات الحركية (مثل منظومات الأسلحة الفتاكة الذاتية التشغيل) لتشمل القدرات غير الحركية التي يمكن أن تكون تصادمية (مثل منظومات الحرب السيبرانية الذاتية التشغيل أو قدرات التشويش المتكيف على الرادارات أو القدرات على شن الحرب الإلكترونية) أو داعمة (مثل اللوجستيات أو الإجلاء الطبي أو المراقبة القتالية) في الأدوار العسكرية. وينبغي لهذه المناقشة أيضا أن تشمل قدرات أخرى قد يكون لها تأثير مباشر على التوازن الاستراتيجي، كتحسين الاستشعار (مثل السواتل أو الاستشعار المضاد للغواصات) أو الاستخبارات أو التخطيط الحربي.

ولا تزال إندونيسيا ملتزمة التزاماً راسخاً بصون السلام والأمن الدوليين على النحو المنصوص عليه في ديباجة الدستور الإندونيسي. إن إندونيسيا، إذ تسترشد بهذا الالتزام، تؤمن بأن استخدام الذكاء الاصطناعي في المجال العسكري يجب أن يكون محكوما بطريقةٍ تعزز السلام والأمن وأهداف التنمية المستدامة. ويجب أن يكون الذكاء الاصطناعي قوة للسلام والأمن لا عاملا دافعا لانعدام الأمن أو النزاع أو الخصومة الاستراتيجية.

ومع أن الذكاء الاصطناعي في حد ذاته ليس سلاحا فإن إندونيسيا تدرك أنه بمثابة مضاعِف للقوة ومضخِّم للتهديد، وقادر على توليد فوائد كبيرة ومخاطر جسيمة للسلام والأمن الدوليين. ويثير استخدام الذكاء الاصطناعي في المجال العسكري العديد من المسائل الأخلاقية والقانونية والمعنوية والتقنية التي ينبغي النظر فيها بعناية والتداول بشأنها لجهة امتثالها للقانون الدولي، بما في ذلك القانون الدولي الإنساني والقانون الدولي لحقوق الإنسان.

فمن ناحية، يُعتقد أن الذكاء الاصطناعي يوفر مجموعة واسعة من الإمكانات. إذ في إمكانه زيادة القدرة على تجهيز البيانات؛ وزيادة الكفاءة والتحديد والدقة التشغيلية؛ وإمكانية تحسين الامتثال للقانون الدولي الإنساني، عبر دعم تقييم التناسُب والتدابير الاحترازية للحد من الضرر الذي قد يلحق بالمدنيين. ويمكنه أيضا أن يعزز القدرات في مجالات الاستخبارات والمراقبة والاستطلاع، وأن يدعم اللوجستيات والتخطيط، ويحسّن إدارة شؤون الأفراد.

ومن ناحية أخرى، يحمل الذكاء الاصطناعي معه مجموعة من المخاطر والعواقب، بينها إمكانية تأجيج سباقات التسلح، وانتشاره إلى جهات فاعلة من غير الدول، وإتاحة إمكان إساءة الاستخدام الإجرامي وغير المسؤول، ومفاقمة الاختلال في القوة العسكرية من خلال التفوق التكنولوجي، وزيادة عدم الاستقرار وسوء التقدير والتصعيد والضبابية القانونية. وتشمل المخاطر التقنية أيضاً مواطن الضعف السيبرانية، وأعطال النظام، وانحياز البيانات، والخطأ في تحديد الهدف، وغير ذلك من أوجه عدم اليقين التشغيلي.

وتشعر إندونيسيا بالقلق بشكل خاص إزاء المخاطر الوجودية الناشئة عن الدمج المحتمل للذكاء الاصطناعي في منظومات القيادة والتحكم والاتصالات النووية. وتؤكد إندونيسيا من جديد موقفها المبدئي بأن استخدام الأسلحة النووية والتهديد باستخدامها ينتهكان القانون الدولي وأننا بحاجة إلى اتخاذ إجراءات عاجلة وحاسمة لدعم وتعزيز القواعد المناهضة للأسلحة النووية. إن اعتماد الذكاء الاصطناعي في منظومات الأسلحة النووية يؤدي إلى تفاقم المخاطر الوجودية لاستخدام الأسلحة النووية، سواء كان ذلك عن قصد أو عن غير قصد أو عن طريق الخطأ، ويزيد من المخاطر النووية. وهذا الأمر يشكل تهديدا لأمن جميع الأمم. لذا تحث إندونيسيا جميع الدول الحائزة للأسلحة النووية على أن تعيد تقييم اعتمادها على الأسلحة النووية وعلى أن تعيد تأكيد التزامنا الجماعي بعالم خالٍ من الأسلحة النووية. وإلى أن تتحقق الإزالة

التامة للأسلحة النووية، يجب على الدول الحائزة للأسلحة النووية أن تحافظ على تحكّم ومسؤولية ومساءلة بشرية حقيقية إزاء الأسلحة النووية ومنظومات إيصالها في سياق تطوير الذكاء الاصطناعي.

وبالنظر إلى هذه الاعتبارات، تحث إندونيسيا على اتباع نهج احترازي في مواجهة تحديات استخدام الذكاء الاصطناعي في المجال العسكري. وتشدد إندونيسيا على أن أوجه تطوير وتطبيق واستخدام الذكاء الاصطناعي في المجال العسكري يجب أن تكون منظمة بما يتيح تسخير فوائده والتخفيف من مخاطره. ويجب على هذه الحوكمة أن تخدم السلام والأمن والازدهار الجماعية لجميع الأمم. وعليه، تطرح إندونيسيا النقاط الرئيسية التالية.

أولاً، تؤكد إندونيسيا على وجوب التمسك بالقانون الدولي طوال دورة حياة تكنولوجيات الذكاء الاصطناعي، ويشمل ذلك ميثاقَ الأمم المتحدة والقانون الدولي الإنساني والقانون الدولي لحقوق الإنسان ومعاهدات نزع السلاح وعدم الانتشار. وينبغي للدول أن تجري مراجعات قانونية في جميع المراحل، بدءاً من الشراء وانتهاءً بالتقييم. ويجب على الدول أن تضمن المساءلة عن تطويرها وتطبيقها للذكاء الاصطناعي في المجال العسكري، بما في ذلك مشروعية تطبيقات الذكاء الاصطناعي في إدارة الحروب أو الأعمال العدائية. وفي ظل غياب هذه القوانين التي تنظم استخدام الذكاء الاصطناعي في المجال العسكري، من المهم التشديد على أنه ينبغي للاستخدام أن يكون محكوما بقوانين الإنسانية وبما يمليه الضمير العام.

وإضافة إلى القانون الدولي، ينبغي للاعتبارات الأخلاقية أن تكون مكمّلة للأطر القانونية في توجيه حوكمة استخدام الذكاء الاصطناعي في المجال العسكري. ويجب الترويج لمبادئ من قبيل التتبع والمساءلة والمسؤولية وقابلية التفسير والإنسانية والشفافية والإنصاف والعدالة في تطوير الذكاء الاصطناعي وتطبيقه.

ثانيا، تشدد إندونيسيا على الدور الأساسي للعنصر البشري في ضمان المساءلة والمسؤولية على جميع المستويات، سواء على مستوى الدولة أو الشركات أو الأفراد، في تصميم وتطوير ونشر واستخدام الذكاء الاصطناعي في المجال العسكري.

ويجب أن يظل تطوير وتطبيق واستخدام الذكاء الاصطناعي في المجال العسكري متمحورا حول الإنسان وأن يكون محكوما بما يخدم مصالح البشرية. ويجب الإبقاء على التحكم البشري الفعال والمجدي وتدعيمه عبر التدريب، لا سيما في القرارات التي تنطوي على استخدام القوة. ويجب أن تنطوي القرارات الحاسمة على التقدير والتدخل والإشراف والتحكم البشرية. وعلاوة على ذلك، ومع أن "التحكم البشري المجدي" يحظى بقبول متزايد في تنظيم استخدام الذكاء الاصطناعي في المجال العسكري، فإن إندونيسيا ترى أن هذا المفهوم لم يجب بعد بشكل شافٍ على الأسئلة القانونية والأخلاقية والتقنية والناظمة المرتبطة بهذا الاستخدام. ويتعين الاتفاق على ما يستلزمه التحكمُ البشري "المجدي" من الناحية العملية.

ومع أن حوكمة الذكاء الاصطناعي ستنظم في المقام الأول سلوك الدول، فإنه يجب عليها أيضا أن تطال الجهات المعنية المدنية، لا سيما شركات التكنولوجيا المنخرطة في استخدام الذكاء الاصطناعي في المجال العسكري. ويجب على الدول أن تضمن امتثال القطاع الخاص للقانون الدولي والمعايير الأخلاقية الدولية وأن تدعم في الوقت عينه نمو منظومة الابتكار في مجال الذكاء الاصطناعي. ويتحمل الباحثون والشركات مسؤولية ضمان أن تكون تكنولوجيا الذكاء الاصطناعي الخاصة بهم موثوقة وآمنة ومأمونة وخاضعة للمساءلة وتحت تحكّم بشري خاضع للمساءلة. كما ينبغي لهم أن يكونوا مسؤولين عن رصد المخاطر التي ينطوي عليها منتجهم والإبلاغ عنها والتصدي لها.

ثالثا، تشدد إندونيسيا على الحاجة الملحة إلى وضع أطر حوكمة قانونية وناظمة متعددة الأطراف وجامعة وشاملة. ويجب أن تراعي هذه الأطر مصالح جميع الدول بغضّ النظر عن مستوى تطور الذكاء الاصطناعي في تلك الدول. ويجب أن يكون لجميع الدول صوت متساوٍ في رسم ملامح القواعد والمعايير الناظمة لاستخدام الذكاء الاصطناعي في المجال العسكري بما يضمن التمثيل العادل ويعزز الثقة العالمية.

إن مشاركة الجهات المعنية على نطاق واسع أمر بالغ الأهمية نظراً إلى الأبعاد الأخلاقية والقانونية والتقنية المتعددة الأوجه للذكاء الاصطناعي. كما أن انخراط التخصصات والثقافات المختلفة ضروري لضمان توافُق منظومات الذكاء الاصطناعي مع القانون الدولي والقانون الإنساني وحقوق الإنسان والالتزامات بنزع السلاح قبل تطبيقها في المجال العسكري.

رابعا، من الأهمية بمكان مواصلة وعي المخاطر والتحديات والآثار الناجمة عن تطوير ونشر واستخدام الذكاء الاصطناعي في المجال العسكري، سواء كانت تكنولوجية أو غير تكنولوجية، وتشجيع إجراء مناقشة مجدية بهذا الشأن. وتُبرز إندونيسيا أهمية إجراء تقييم مستمر للآثار الأوسع نطاقاً للذكاء الاصطناعي العسكري على السلام والأمن الدوليين، لا سيما في سياق عدم الانتشار ونزع السلاح. ومن الضروري إجراء مزيد من الدراسات الشاملة لتكوين فهمٍ لهذه التأثيرات التي لا تزال غير مستكشفة بشكل كافٍ.

إن تحديد المخاطر المرتبطة بتطوير ونشر واستخدام الذكاء الاصطناعي في المجال العسكري سيدعم التنبؤ القائم على الأدلة وتقييم المخاطر ووضع تدابير التخفيف من المخاطر في نهاية المطاف.

ومن المهم أيضا تعزيز فهم المخاطر المرتبطة باستخدام الذكاء الاصطناعي في المجال العسكري والتوعية بها. وفي هذا الصدد، ينبغي تعزيز الشفافية من خلال جملة أمور منها الاطلاع المتبادل على السياسات والاستراتيجية الوطنية، لا سيما لتحديد المخاطر وتقييمها والتخفيف منها؛ والاطلاع المتبادل على قدرات الذكاء الاصطناعي في المجال العسكري، عند الاقتضاء، لزيادة المساءلة وتدابير بناء الثقة؛ وتبادل الدروس المستفادة وأفضل الممارسات عبر الحدود والصناعات والقطاعات.

خامسا، يجب ألا تعوق حوكمةُ الذكاء الاصطناعي التطورَ التكنولوجي أو تحد من وصول الدول النامية إلى الذكاء الاصطناعي. وينبغي للأطر أن تتجنب فرض شروط أو حواجز تقيد الوصول العادل. ومن الضروري اتّباع نهج متوازن يتصدى لمخاطر من قبيل الانتشار ويضمن في الوقت عينه للدول ذات الموارد المحدودة إمكانية الوصول إلى الذكاء الاصطناعي.

وأخيراً، يجب أن تركز حوكمة الذكاء الاصطناعي بقوةٍ على سد الفجوة الرقمية وفي مجال الذكاء الاصطناعي. فالبلدان النامية تواجه قيودا كبيرة، لا في ما يتعلق بقدرات الذكاء الاصطناعي فحسب بل أيضا في قدرتها على إدارة هذه التكنولوجيا بفعالية. وإذا ظلت هذه الفجوة قائمة، فإن جهود الحوكمة العالمية ستتقوض إذ لا يزال العديد من الدول غير مجهز بما فيه الكفاية لمواجهة التحديات المعقدة والعابرة للحدود التي يطرحها الذكاء الاصطناعي.

وتشدد إندونيسيا على الحاجة الملحة إلى معالجة الفجوات الصارخة في المجال الرقمي وفي الذكاء الاصطناعي بين الدول وداخلها، لا سيما في ما يتعلق بالوصول إلى الموارد المالية والبشرية والتقنية. فوجود هذه الفجوات يهدد بتعميق أوجه عدم المساواة العالمية وبزيادة احتمال نشوب النزاعات.

إن السلام والأمن، باعتبارهما من المنافع العامة العالمية، يتطلبان التعاون الدولي بين جميع البلدان، سواء كانت متقدمة النمو أو نامية، لمواجهة التحديات المشتركة والاستفادة من الفوائد الجماعية، بما فيها تلك المتعلقة بتطوير وتطبيق واستخدام الذكاء الاصطناعي في المجال العسكري. وفي هذا السياق، تدعو إندونيسيا إلى تعزيز التعاون والمساعدة الدوليين وتحقيق التوازن بينهما لتعزيز قدرات الذكاء الاصطناعي العالمية وأطر حوكمته. ويجب تحقيق هذا التعاون على أساس منصف ومتفق عليه، مع مراعاة الاحتياجات والسياقات المحددة للبلدان النامية. ويشمل ذلك على سبيل المثال لا الحصر مبادرات في مجال بناء القدرات والتعليم ونقل التكنولوجيا والتعلم مدى الحياة والتدريب التقني والبحوث المشتركة وتبادل المعارف.

ويجب أن يكون هذا التعاون متعدد المستويات لا فقط بين الدول والمنظمات الدولية بل أيضا عبر القطاعات داخل البلدان. وينبغي تشجيع الشراكات بين القطاعين العام والخاص من أجل تعزيز الابتكار المسؤول والتوعية داخل القطاعات المعنية بالآثار التي قد تترتب على تكنولوجياتها في السلام والأمن الدوليين.

إن التعاون الدولي يتسم بأهمية حيوية لا فقط لإيجاد حلول للفجوات سواء كانت رقمية أو في مجال الذكاء الاصطناعي، بل أيضا لإشاعة بيئة مؤاتية لبناء الثقة بين الدول. ويمكنه أن يساعد في الحد من الانقسام الجيوسياسي والمنافسة في مجال الذكاء الاصطناعي. ويجب أن يكون التعاون الدولي مرتكزا على مبادئ المساواة والثقة والمنفعة المتبادلة واحترام السيادة والتضامن لتمهيد الطريق أمام تعاون مجدٍ يشمل نقل التكنولوجيا وتبادل المعارف.

إن إندونيسيا تدرك أيضاً قيمة تدعيم آليات التعاون الإقليمي التي تراعي الخصوصيات المحلية والإقليمية. ويمكن أن تكون هذه الآليات بمثابة لبِنات بناءٍ أساسية نحو توافق عالمي أوسع نطاقا يوفر في الوقت عينه مساحة لمداولات أكثر تفصيلا ومراعاة للسياق.

إيران (جمهورية - الإسلامية)

[الأصل: بالإنكليزية]
[12 آذار/مارس 2025]

استجابةً للطلب الموجه إلى الأمين العام بموجب الفقرة 7 من القرار 239/79، الذي التمست فيه الجمعية العامة آراء الدول الأعضاء بشأن الفرص والتحديات التي يطرحها تطبيق الذكاء الاصطناعي في المجال العسكري على السلام والأمن الدوليين مع التركيز بشكل خاص على مجالات أخرى غير الأسلحة الفتاكة الذاتية التشغيل، تقدم جمهورية إيران الإسلامية بموجب هذه الوثيقة آراءها بهذا الشأن.

لقد بات الذكاء الاصطناعي أحد المحركات الرئيسية للتغيير في عالم اليوم، فهو ترك بصمة لا تُمحى على كيفية تطور الصناعة العسكرية في المستقبل القريب، مما يؤثر على السلام والأمن الدوليين في الصميم. وتجهد الجهات الفاعلة الحكومية وغير الحكومية للدفع بخططها المتنافسة في مجال الذكاء الاصطناعي والتي لا يمكن تركها من دون تنظيم. وبالنظر إلى الدور القيادي للجهات الفاعلة من غير الدول، وضرورة تحقيق توازن بين الإجراءات الناظمة والمبتكرة والاتجاهات المسجّلة، من الحيوي أن تظل السلطة الناظمة حكراً على الصلاحية السيادية للدول الأعضاء.

ومن منظور موضوعي، كما هو الحال بالنسبة إلى التكنولوجيات الأخرى المستخدمة في الفضاء السيبراني والفضاء الخارجي، تؤيد جمهورية إيران الإسلامية التطبيق السلمي حصرا للذكاء الاصطناعي، مع الأخذ في الاعتبار أنه في ظل الظروف المناسبة، يمكن للكيانات العسكرية أيضا أن تستفيد من مكاسب الذكاء الاصطناعي بشكل سلمي.

ونظراً إلى تفاوت مستويات التنمية بين الدول، من الأهمية بمكان ضمان عدم تطور الفجوة الرقمية إلى فجوة في الذكاء الاصطناعي. ولا يمكن ضمان الطابع الشامل للجميع للإجراءات الناظمة المتعلقة بالذكاء الاصطناعي إلا ضمن الإطار التوافقي الذي توفره الأمم المتحدة. واتّباع هذا النهج يحافظ على سيادة الدول الأعضاء، ويشيع بيئة من التطوير المنصف للذكاء الاصطناعي للجميع، ويوفر مرونة مبتكرة تتيح ازدهار قطاع الذكاء الاصطناعي. والدور المركزي للأمم المتحدة في المسائل الناظمة المتعلقة بالذكاء الاصطناعي يتصدى للطابع الحصري الذي تتسم به النهج الوطنية إزاء هذه المسألة. وعليه، يجب أن يكون الطابع الشامل والنهج التوافقي هما المبدأ الأعلى الذي يحكم هذه المسألة ذات الأهمية الحيوية.

ورغم المناقشات الجارية حول الذكاء الاصطناعي في المحافل الدولية المختلفة، فإن فهمنا لهذه القضية ولتبعاتها على السلام والأمن الدوليين لا يزال غير مكتمل. ومن السابق لأوانه تأكيد الانطباق الكامل للقانون الدولي الإنساني والقانون الدولي لحقوق الإنسان على الذكاء الاصطناعي. وإزاء جسامة هذه الظاهرة الجديدة والمتطورة بسرعة قد يحتاج الإطار القانوني الدولي إلى تكييف وتطوير فريدين.

وفي ما يتعلق بالجهود الناظمة على الصعيد الدولي، تؤيد جمهورية إيران الإسلامية وضع ترتيبات ملزمة قانوناً بين الدول الأعضاء باعتبارها مسار العمل المفضل لديها، عوض تحديد معايير أو وضع صكوك سياسية.

إن جمهورية إيران الإسلامية، انسجاما مع موقفها المبدئي من نزع السلاح، ترفض أي نهج تمييزي أو مشروط أو ذي دوافع سياسية أو أي ازدواجية في المعايير. ويجب على المصطلحات التي تستخدمها الجمعية العامة أن تعكس تاليا جوا من الوحدة والتوافق. ومن هذا المنطلق، فإن مفاهيم من قبيل "التطبيق المسؤول" هي مفاهيم مجردة للغاية لا يمكن أن تنظم مجالا يستلزم تعريفَه بوضوح ودقة. وبالتالي، فإن من شأن مفهوم مجرد كهذا أن يفسح في المجال أمام إساءة التفسير وأن يفتح الباب أمام اتّباع نهج مسيّس. وعليه، تعرب جمهورية إيران الإسلامية عن اعتراضها الشديد على استخدام مثل هذه المصطلحات غير الموضوعية. وهي تؤيد وتقترح الاستعاضة عن مصطلح "التطبيق المسؤول" بمصطلح "التطبيق السلمي" في أي صك مقبل.

إسرائيل

[الأصل: بالإنكليزية]
[10 نيسان/أبريل 2025]

تحيط إسرائيل علما باتخاذ قرار الجمعية العامة 239/79، ووفقا للفقرة 7 من القرار، تتشرف بأن تقدم مساهمتها الوطنية في تقرير الأمين العام إلى الجمعية في دورتها الثمانين لكي تواصل الدول الأعضاء مناقشتها.

وترى إسرائيل أن مفهوم الذكاء الاصطناعي يحتمل حاليًا مجموعة من التفسيرات الممكنة التي قد يتم تنقيحها بمرور الوقت.

ومن الواضح أن استخدام الذكاء الاصطناعي في المجال العسكري أصبح أكثر شيوعاً وتواتراً من أي وقت مضى. وقد صوتت إسرائيل لصالح قرار الجمعية العامة المذكور أعلاه، وهي تشجع الدول وجميع أصحاب المصلحة على المشاركة في مناقشة تأخذ في الحسبان الاعتبارات المشروعة لجميع الدول، بما في ذلك الاعتبارات الأمنية والإنسانية والاقتصادية والتنموية، وذلك مع الحفاظ على طابع مهني وغير مُسيَّس.

ومن أجل إجراء مناقشة جادة ومسؤولة حول الذكاء الاصطناعي في المجال العسكري والتي قد يكون لها تأثير مفيد في نهاية المطاف، نعتقد أنه يجب اعتماد نهج عملي ومتوازن وتدريجي.

ونظراً لأن التكنولوجيا تجلب مجموعة واسعة من الفرص في كل مجال تقريباً، بما في ذلك المجال العسكري، فإننا نرحب باستكشاف الفوائد التي يمكن أن تجلبها هذه التطورات وسبل تحقيقها، فضلاً عن المخاطر المحتملة وسبل التخفيف من حدتها. وترى إسرائيل أن التكنولوجيات الناشئة، مثل تكنولوجيات الذكاء الاصطناعي، قد تعمل أيضًا على تعزيز الالتزام بالقانون الدولي الإنساني القائم. وتقتضي هذه الفرص المحتملة عدم وصم هذه التكنولوجيات بشكل سلبي.

وتظل إسرائيل صوتًا بنّاءً في الخطاب العالمي حول استخدام الذكاء الاصطناعي في الأغراض العسكرية. وقد أقرت مؤخراً الإعلان السياسي بشأن الاستخدام العسكري المسؤول للذكاء الاصطناعي وخاصية التشغيل الذاتي، وذلك بقيادة الولايات المتحدة. ونحن نتطلع إلى الانضمام إلى الاجتماعات المستقبلية التي تتناول هذه المبادرة ومواصلة تعزيز الاستخدام العسكري المسؤول للذكاء الاصطناعي وخاصية التشغيل الذاتي.

وكجزء من الإعلان، وكذلك في سياقات أخرى، فكرت الدول في السنوات الأخيرة في وضع توجيهات بشأن تطوير الذكاء الاصطناعي واستخدامه في المجال العسكري، سواء على الصعيد المحلي أو الدولي. ويبدو أن بعض المبادئ الأساسية والأكثر شيوعا ضمن هذه التوجيهات، والتي قد تكون ذات صلة أيضا بالمناقشات في سياق القرار 239/79، تتمثل في ما يلي:

- يجب أن يكون الاستخدام العسكري للذكاء الاصطناعي متوافقاً مع القانون الدولي الساري.
- يجب أن يكون مسؤولا وأن يعزز الأمن الدولي.
- يجب على الدول ضمان المساءلة فيما يتعلق باستخدام قدرات الذكاء الاصطناعي وفقاً للقانون الدولي الساري، بما في ذلك تشغيلها في إطار سلسلة من القيادة والتحكم البشريين تتسم بالمسؤولية.

ومن بين التدابير العملية التي ينبغي للدول اتخاذها لتفعيل هذه المبادئ ما يلي:

- ينبغي للدول اتخاذ الخطوات المناسبة، مثل استعراض القوانين، لضمان استخدام قدرات الذكاء الاصطناعي في المجال العسكري بما يتوافق مع التزاماتها بموجب القانون الدولي، ولا سيما القانون الدولي الإنساني.

- ينبغي للدول اتخاذ التدابير المناسبة لضمان التطوير المسؤول لقدرات الذكاء الاصطناعي في المجال العسكري ونشرها واستخدامها. ينبغي تنفيذ هذه التدابير في المراحل ذات الصلة طوال دورة حياة قدرات الذكاء الاصطناعي في المجال العسكري.
- ينبغي للموظفين المعنيين توخي الحذر المناسب عند تطوير ونشر واستخدام قدرات الذكاء الاصطناعي في المجال العسكري، بما في ذلك منظومات الأسلحة التي تتضمن هذه القدرات.
- نبغي أن يشرف كبار المسؤولين بشكل فعال ومناسب على تطوير ونشر قدرات الذكاء الاصطناعي في المجال العسكري ذات التطبيقات الشديدة العواقب، بما في ذلك، على سبيل المثال لا الحصر، منظومات الأسلحة التي تنطوي على مثل هذه القدرات.
- ينبغي للدول دعم الجهود المناسبة لضمان استخدام قدرات الذكاء الاصطناعي في المجال العسكري بشكل مسؤول وقانوني، ومتابعة المناقشات المستمرة مع الدول الأخرى حول كيفية نشر قدرات الذكاء الاصطناعي في المجال العسكري واستخدامها على هذا النحو.

وتقدِّر إسرائيل إجراء مناقشات شاملة متعددة الأطراف حول الذكاء الاصطناعي في المجال العسكري، وآثاره على الأمن الدولي، من شأنها أن تحقق التوازن الصحيح بين الضرورة العسكرية والاعتبارات الإنسانية.

إيطاليا

[الأصل: بالإنكليزية]
[11 نيسان/أبريل 2025]

الرئاسة الإيطالية لمجموعة الدول السبع

وُضع الذكاء الاصطناعي في صلب المناقشات السياسية والتقنية طوال فترة الرئاسة الإيطالية لمجموعة الدول السبع في عام 2024. وأقرت قمة القادة التي عُقدت في أبوليا بتأثير الذكاء الاصطناعي على المجال العسكري والحاجة إلى إطار عمل لتطويره واستخدامه بشكل مسؤول.

وفي الفترة من 18 إلى 20 تشرين الأول/أكتوبر 2024، عُقد أول اجتماع وزاري لمجموعة الدول السبع بشأن الدفاع في نابولي. وفي تلك المناسبة، أكد وزراء دفاع مجموعة الدول السبع من جديد عزمهم على مواجهة التحديات الأمنية بطريقة متماسكة وملموسة، في وقت يتسم بقدر كبير من عدم الاستقرار. كما شددوا على ضرورة اتباع نهج أكثر تعاوناً في مجال البحث والتطوير المتعلق بالدفاع، بما في ذلك من حيث تبادل الخبرات والمعارف والاستفادة منها، مع تعزيز بيئة آمنة لمنع الوصول إليها لأغراض خبيثة، من أجل الحفاظ على الميزة التنافسية، بما في ذلك في مجال التكنولوجيات الناشئة والكاسحة.

وأخيرًا، أقر فريق المديرين المعني بعدم الانتشار التابع لمجموعة الدول السبع في بيانه بالتأثير العميق للتكنولوجيات الكاسحة الناشئة، مثل الذكاء الاصطناعي، على الحد من التسلح وعدم الانتشار ونزع السلاح، وكذلك على مستقبل العمليات العسكرية.

أولا - عملية الذكاء الاصطناعي المسؤول في المجال العسكري

تقدر إيطاليا عملية الذكاء الاصطناعي المسؤول في المجال العسكري التي أطلقتها هولندا وجمهورية كوريا في عام 2023 بهدف توفير منصة لمناقشة الفرص والتحديات والمخاطر الرئيسية المرتبطة بالتطبيقات العسكرية للذكاء الاصطناعي. وفي القمة الثانية لتسخير الذكاء الاصطناعي المسؤول في المجال العسكري، التي عُقدت في سول في عام 2024، أقرت إيطاليا مخطط العمل، وهو وثيقة تحدد المبادئ الرئيسية للحوكمة المسؤولة للذكاء الاصطناعي، بما في ذلك أهمية الامتثال للقانون الدولي، والمسؤولية الإنسانية والمساءلة، وموثوقية أنظمة الذكاء الاصطناعي وجدارتها بالثقة، والمشاركة البشرية المناسبة في تطوير الذكاء الاصطناعي ونشره واستخدامه في المجال العسكري.

وتشدد الدول التي أقرت المخطط على ضرورة منع استخدام تكنولوجيات الذكاء الاصطناعي للمساهمة في انتشار أسلحة الدمار الشامل، وأهمية عدم تقويض جهود الحد من التسلح ونزع السلاح وعدم الانتشار. وعلاوة على ذلك، ومن أجل بلورة فهم مشترك لتكنولوجيا الذكاء الاصطناعي وتطبيقاتها في المجال العسكري، يدعو المخطط الدول إلى الالتزام بإجراء المزيد من المناقشات ووضع إجراءات فعالة لاستعراض القوانين وتدابير بناء الثقة والتدابير المناسبة للحد من المخاطر. وفي هذا الإطار، يعد تبادل المعلومات والممارسات الجيدة، فضلاً عن المشاركة الفعالة لأصحاب المصلحة الآخرين، أمرًا حاسمًا لإحراز تقدم في هذا النقاش.

ثانيا - الإعلان السياسي بشأن الاستخدام العسكري المسؤول للذكاء الاصطناعي وخاصية التشغيل الذاتي

تقدِّر إيطاليا أيضًا الإعلان السياسي بشأن الاستخدام العسكري المسؤول للذكاء الاصطناعي وخاصية التشغيل الذاتي. وتأكيدًا على أن الاستخدام العسكري للذكاء الاصطناعي يمكن وينبغي أن يكون أخلاقيًا ومسؤولًا ومعزِّزًا للأمن الدولي، أقرت الدول المصدقة على ضرورة تنفيذ مجموعة من التدابير في سبيل تطوير قدرات الذكاء الاصطناعي في المجال العسكري ونشرها واستخدامها. وتلتزم الدول على وجه الخصوص بما يلي: تقليل التحيز غير المقصود فيما يخص قدرات الذكاء الاصطناعي في المجال العسكري؛ والتأكد من خضوع سلامتها وأمنها وفعاليتها خضعت لاختبارات مناسبة وصارمة؛ وتنفيذ الضمانات المناسبة لاكتشاف وتجنب العواقب غير المقصودة والاستجابة بفعالية في مثل هذه الحالات. وعلاوة على ذلك، من المهم أن يتم تحديد سلسلة من القيادة والتحكم البشريين تتسم بالمسؤولية وأن يتم استخدام قدرات الذكاء الاصطناعي في المجال العسكري بطريقة تتفق مع الالتزامات الدولية.

ثالثا - ميثاق المستقبل

في أيلول/سبتمبر 2024، اعتمد قادة العالم ميثاق المستقبل الذي يؤكد من جديد التزاماتهم العالمية ويمكّن الدول من مواجهة التحديات والفرص الجديدة والناشئة. ففي إطار الإجراء 27، تُشجَّع الدول على اغتنام الفرص المرتبطة بالتكنولوجيات الناشئة، بما في ذلك الذكاء الاصطناعي، وفي الوقت نفسه معالجة المخاطر المحتملة التي يشكلها سوء استخدامها. وعلى وجه الخصوص، ستواصل الدول الأعضاء تقييم هذه المخاطر في التطبيقات العسكرية للذكاء الاصطناعي والفرص الممكنة طوال دورتها العُمرية، بالتشاور مع الجهات المعنية صاحبة المصلحة.

رابعا – إعلان باريس بشأن الحفاظ على التحكم البشري في منظومات الأسلحة القائمة على الذكاء الاصطناعي

صادقت إيطاليا أيضا في الآونة الأخيرة على إعلان باريس بشأن الحفاظ على التحكم البشري في منظومات الأسلحة القائمة على الذكاء الاصطناعي، والذي تم اعتماده على هامش قمة العمل في مجال الذكاء الاصطناعي التي عُقدت في باريس في الفترة من 6 إلى 11 شباط/فبراير 2025. وتأكيداً على أنه لا يمكن أبداً نقل المسؤولية والمساءلة إلى الآلات، التزمت الدول المصدقة على نهج يركز على الإنسان في تطوير تطبيقات الذكاء الاصطناعي ونشرها واستخدامها في المجال العسكري. كما التزمت بضمان أن يكون نشر الذكاء الاصطناعي في القطاع العسكري متوافقًا تمامًا مع القانون الدولي والقانون الدولي الإنساني، مع تعزيز البحث والتطوير والابتكار في مجال تكنولوجيا الذكاء الاصطناعي.

خامسا – فريق الخبراء الحكوميين المعني بالتكنولوجيات الناشئة في مجال منظومات الأسلحة الفتاكة الذاتية التشغيل

تترتب أيضا على التطورات السريعة في الذكاء الاصطناعي والتعلم الآلي آثار كبيرة فيما يخص دور خاصية التشغيل الذاتي في منظومة الأسلحة. وترى إيطاليا أن اتفاقية حظر أو تقييد استعمال أسلحة تقليدية معينة يمكن اعتبارها مفرطة الضرر أو عشوائية الأثر، التي تجمع بين الخبرة الدبلوماسية والقانونية والعسكرية لممثلي الحكومات والمنظمات الدولية والمؤسسات المتخصصة، هي إلى حد بعيد أنسب محفل لمعالجة المسائل الراهنة والناشئة المتعلقة بتطوير منظومات الأسلحة واستخدامها. وتساهم إيطاليا بنشاط في المناقشات التي تجري في إطار فريق الخبراء الحكوميين المعني بالتكنولوجيات الناشئة في مجال منظومات الأسلحة الفتاكة الذاتية التشغيل، الذي أُطلق تحت رعاية الاتفاقية، وهي ملتزمة بالمضي قدماً في المناقشات بشأن وضع عناصر صك يبرم في المستقبل وفقاً للولاية المتفق عليها في اجتماع الأطراف المتعاقدة السامية في الاتفاقية لعام 2023.

وترى إيطاليا أن هذا الصك يجب أن ينص على محظورات وقواعد تنظيمية واضحة، بحيث يتم اعتماده في نهاية المطاف كبروتوكول إضافي للاتفاقية. ووفقا لهذا النهج، فإن منظومات الأسلحة الفتاكة الذاتية التشغيل التي لا يمكن تطويرها واستخدامها وفقا للقانون الدولي الإنساني ستكون محظورة بحكم الواقع. ومن ناحية أخرى، ستخضع للتنظيم المنظومات التي تتسم باستقلالية اتخاذ القرار في الوظائف الحاسمة، والتي يمكن تطويرها واستخدامها في امتثال كامل للقانون الدولي الإنساني. وترى إيطاليا أن العنصر البشري في الواقع بالغ الأهمية بالنسبة للدورة العمرية الكاملة لمنظومات الأسلحة الفتاكة الذاتية التشغيل، أي التصميم والاستحداث والإنتاج والنشر والاستخدام. كما ينبغي الإبقاء على مستوى مناسب من التقدير والتحكم البشريين لضمان المسؤولية والمساءلة بموجب القانون الدولي الإنساني.

اليابان

[الأصل: بالإنكليزية]
[11 نيسان/أبريل 2025]

طلبت الجمعية العامة، في قرارها 239/79، إلى الأمين العام أن يلتمس آراء الدول الأعضاء والدول المراقبة بشأن الفرص والتحديات التي يطرحها تطبيق الذكاء الاصطناعي في المجال العسكري على السلام والأمن الدوليين، مع التركيز بوجه خاص على مجالات غير منظومات الأسلحة الفتاكة الذاتية

التشغيل، وأن يقدم تقريرا موضوعيا يلخص تلك الآراء ويصنف المقترحات المعيارية القائمة والناشئة، ويحتوي على مرفق يتضمن تلك الآراء، إلى الجمعية العامة في دورتها الثمانين لكي تواصل الدول مناقشتها. وتقدم اليابان بموجب هذه الوثيقة آراءها بشأن هذا الموضوع بغرض المساهمة في إعداد التقرير وفي مواصلة النقاش حول هذا الموضوع.

أولا - وجهات نظر عامة

إن اليابان ملتزمة بالحفاظ على نظام دولي حر ومفتوح قائم على سيادة القانون وتعزيزه حتى يتسنى لجميع الناس التمتع بالسلام والاستقرار والازدهار، وبتعزيز الدبلوماسية لتحقيق عالم آمن ومأمون تصان فيه كرامة الإنسان. وتماشيا مع هذه الأهداف، شاركت اليابان بنشاط في الجهود الرامية إلى تعزيز السلام والأمن الدوليين وكذلك تحديد الأسلحة ونزع السلاح.

وترى اليابان أنه ينبغي دراسة تطبيق الذكاء الاصطناعي في المجال العسكري بطريقة شاملة، مع فهم كافٍ لمخاطره وفوائده، ومراعاة الاعتبارات الإنسانية والمنظورات الأمنية على حد سواء. ومن المفيد تعزيز فهم تطبيق الذكاء الاصطناعي في المجال العسكري، وتعزيز الجهود الواقعية والعملية الرامية إلى استخدامه على نحو مسؤول من أجل تعظيم فوائده مع الحد من مخاطره.

وفيما يتعلق بتطبيق الذكاء الاصطناعي في المجال العسكري، تؤيد اليابان الرأي القائل، أولاً، بأن القانون الدولي الحالي ينطبق على ما يطرأ على مدى الدورة العُمرية للذكاء الاصطناعي من مسائل تخضع لأحكامه؛ وثانيًا، بأنه يجب تطبيق قدرات الذكاء الاصطناعي بطريقة مسؤولة؛ وثالثًا، بأنه يبقى الإنسان مسؤولاً وخاضعاً للمساءلة عن استخدامها وآثارها. وتؤكد اليابان أيضًا على الحاجة إلى تعزيز الشفافية كتدبير مهم لبناء الثقة من أجل تعظيم الفوائد مع الحد من المخاطر.

ثانيا - آراء اليابان ونهجها فيما يتعلق بالفرص والتحديات التي تواجه السلام والأمن الدوليين بسبب تطبيق الذكاء الاصطناعي في المجال العسكري

الفرص المتاحة

الآراء

تؤدي التطورات السريعة في العلوم والتكنولوجيا، بما في ذلك الذكاء الاصطناعي، إلى تغيير جذري في نموذج الأمن. وتسعى الدول جاهدة لاستحداث تكنولوجيات متطورة يمكن أن تغير طابع الحرب بشكل كبير وبالتالي يمكن أن تغير قواعد اللعبة، وأصبح من الصعب للغاية من الناحية العملية التمييز بين التكنولوجيات المخصصة للاستخدام المدني وتلك المخصصة للأغراض الأمنية. وينطوي الذكاء الاصطناعي على إمكانات غير عادية لتحويل كل جانب من جوانب الشؤون العسكرية، بما في ذلك العمليات العسكرية؛ والقيادة والتحكم؛ وأنشطة الاستخبارات والمراقبة والاستطلاع؛ والتدريب؛ وإدارة المعلومات؛ والدعم اللوجستي. وبالنظر إلى الاستخدامات المتنوعة للذكاء الاصطناعي في المجال العسكري، فإن تطبيق الذكاء الاصطناعي قد يحقق فوائد مثل تحسين الدقة والصحة والكفاءة؛ وتعزيز الإلمام بالحالة وفهمها؛ وتسهيل التحليل السريع للمعلومات؛ والحد من الأخطاء البشرية؛ والاقتصاد في اليد العاملة. ويمكن أن يسهم تطبيقه بشكل صحيح في تحسين حماية المدنيين في النزاعات وبناء السلام بعد انتهاء النزاع.

النهج الياباني تجاه الاستفادة من "الفرص"

عند تطبيق الذكاء الاصطناعي في المجال العسكري، لا بد من النظر في ما إذا كان هذا التطبيق فعالاً في التغلب على المشكلات التي يحددها البشر، مع مراعاة وظائف الذكاء الاصطناعي وأوجه قصوره. ولا ينبغي أن يكون تطبيق الذكاء الاصطناعي في حد ذاته هو الهدف، ولا ينبغي النظر إليه بمعزل عن وظائفه وأوجه قصوره. ولذلك، ينبغي للدول أن تضمن أن تكون لقدرات الذكاء الاصطناعي في المجال العسكري استخدامات واضحة ومحددة بشكل جيد وأن تكون مصممة ومهندسة لأداء الوظائف المنشودة. ومع أخذ ذلك في الاعتبار، من المهم تشجيع التوصل إلى فهم دولي مشترك للذكاء الاصطناعي ووظائفه وأوجه قصوره في المجال العسكري، بالإضافة إلى فهم مشترك للتطبيق المحتمل للذكاء الاصطناعي في المجال العسكري. أما فيما يتعلق بتطبيق الذكاء الاصطناعي من قبل سلطات الدفاع، فقد نشرت وزارة الدفاع اليابانية السياسة الأساسية لوزارة الدفاع بشأن تعزيز استخدام الذكاء الاصطناعي في تموز/يوليه 2024، وهي السياسة التي بيَّنت أفكارها الحالية حول وظائف الذكاء الاصطناعي وأوجه قصوره في المجال العسكري ومجالات تطبيق الذكاء الاصطناعي التي تعطيها الأولوية. وفي هذه السياسة الأساسية، وفي ضوء قدرات الذكاء الاصطناعي وأوجه قصوره في الوقت الراهن، حددت وزارة الدفاع المجالات السبعة التالية التي تركز فيها على تطبيق الذكاء الاصطناعي:

- كشف الأهداف وتحديدها
- جمع المعلومات الاستخباراتية وتحليلها
- القيادة والتحكم
- عمليات الدعم اللوجستي
- الأصول غير المأهولة
- الأمن السيبراني
- زيادة كفاءة الأعمال الإدارية

وتشير السياسة الأساسية أيضًا إلى أنه من الضروري أن نضع في اعتبارنا أن الذكاء الاصطناعي يُطبق لدعم اتخاذ القرار البشري، وأن مشاركة البشر ضرورية عند تطبيق الذكاء الاصطناعي.

التحديات

الآراء

يمكن أن ينطوي تطبيق الذكاء الاصطناعي في المجال العسكري على مخاطر إساءة الاستخدام أو الاستخدام الخبيث، والتصعيد وخفض عتبة نشوب النزاعات، وهو ما قد ينشأ عن التحيز والعواقب غير المقصودة وعوامل أخرى. وفي هذا الصدد، تشدد اليابان على ضرورة منع استخدام الذكاء الاصطناعي للمساهمة في انتشار أسلحة الدمار الشامل من قبل الدول والجهات الفاعلة من غير الدول، وتؤكد على أن الذكاء الاصطناعي ينبغي أن يدعم جهود نزع السلاح وتحديد الأسلحة وعدم الانتشار لا أن يعرقلها.

النهج الياباني في مواجهة "التحديات"

في ضوء مخاطر مثل التحيز وسوء الاستخدام والاستخدام الخبيث، ستعمل وزارة الدفاع اليابانية على الحد من المخاطر التي يشكلها الذكاء الاصطناعي، متخذةً كمرجع مفاهيم الذكاء الاصطناعي المتمحور حول الإنسان والسلامة والإنصاف وحماية الخصوصية وضمان الأمن والشفافية والمساءلة، على النحو المبين في المبادئ التوجيهية للذكاء الاصطناعي لأغراض الأعمال التجارية في اليابان المنشورة في نيسان/أبريل 2024، مع الاهتمام أيضًا بالمناقشات الجارية في المجتمع الدولي ومع سلطات الدفاع في بلدان أخرى.

وبالإضافة إلى ذلك، تتابع اليابان عن كثب التأثير المحتمل للتكنولوجيات الناشئة مثل الذكاء الاصطناعي على نزع السلاح النووي وعدم الانتشار النووي. وفي هذا الصدد، ترحب اليابان بالالتزام الذي تعهدت به الولايات المتحدة والمملكة المتحدة وفرنسا في مؤتمر الأطراف في معاهدة عدم انتشار الأسلحة النووية لاستعراض المعاهدة عام 2022، بالحفاظ على التحكم البشري والدور البشري في جميع الإجراءات الحاسمة في إرشاد وتنفيذ القرارات السيادية المتعلّقة باستخدام الأسلحة النووية، وتدعو الدول الأخرى الحائزة للأسلحة النووية إلى أن تحذو حذوها. وعلاوة على ذلك، شدد الفريق الدولي للشخصيات البارزة من أجل إقامة عالم خال من الأسلحة النووية، في توصيته إلى مؤتمر الاستعراض لعام 2026، على ضرورة العمل معًا لمواجهة التحديات والفرص المرتبطة بالتكنولوجيات الناشئة.

ثالثا - آراء حول مستقبل المناقشات والتعاون الدولي

من الضروري اتباع نهج مرن ومتوازن وواقعي لحوكمة الذكاء الاصطناعي في المجال العسكري من أجل مواكبة التطور والتقدم السريع للتكنولوجيات. وتشدد اليابان على أن جهود الذكاء الاصطناعي المسؤول في المجال العسكري يمكن أن تُبذل بالتوازي مع جهود البحث والتطوير والتجريب والابتكار في مجال تكنولوجيا الذكاء الاصطناعي، وهي لا تعيقها.

وتجدر الإشارة إلى أن المناقشات التي تثير إشكالية حول تكنولوجيات محددة للذكاء الاصطناعي قد تؤدي إلى إعاقة التطور التكنولوجي والابتكار في القطاع المدني، مع إمكانية إحداث أثر مثبط. وعلاوة على ذلك، ينبغي مناقشة تطبيق الذكاء الاصطناعي في المجال العسكري بطريقة شاملة للجميع وبمشاركة أصحاب المصلحة وفي ظل تبادل الآراء فيما بينهم.

وفي ضوء الاعتبارات المذكورة أعلاه، تؤيد اليابان بقوة نتائج مؤتمرات القمة الخاصة بتسخير الذكاء الاصطناعي المسؤول في المجال العسكري والإعلان السياسي بشأن الاستخدام العسكري المسؤول للذكاء الاصطناعي وخاصية التشغيل الذاتي وتتوقع انضمام المزيد من الدول إلى هذه المبادرات.

وفيما يتعلق بمنظومات الأسلحة الفتاكة الذاتية التشغيل، تجدر الإشارة إلى أن اليابان تؤيد بقوة استمرار المناقشات في إطار اتفاقية حظر أو تقييد استعمال أسلحة تقليدية معينة يمكن اعتبارها مفرطة الضرر أو عشوائية الأثر، وتتوقع أن تكون المناقشات بشأن تطبيق الذكاء الاصطناعي في المجال العسكري مكمِّلةً ومعزِّزةً للمناقشات الدائرة على مستوى فريق الخبراء الحكوميين المعني بالتكنولوجيات الناشئة في مجال منظومات الأسلحة الفتاكة الذاتية التشغيل المنشأ بموجب الاتفاقية.

وتدرك اليابان أن الشفافية في تطبيق الذكاء الاصطناعي في المجال العسكري أمر مهم كتدبير لبناء الثقة يؤدي إلى الحد من المخاطر، فضلاً عن التعاون الفعال والتعاون فيما بين البلدان. وتدرك اليابان أيضاً أهمية بناء القدرات لتيسير اتباع النهج المسؤول إزاء تطوير ونشر واستخدام الذكاء الاصطناعي في المجال العسكري، وتلتزم بتعزيز التعاون الدولي في مجال بناء القدرات بهدف تقليص الفجوة المعرفية المتعلقة بهذا النهج. وفي هذا الصدد، ستكون أساليب مثل تبادل الممارسات الجيدة والدروس المستفادة مفيدة، وستستفيد اليابان من فرص تبادل وجهات النظر مع بلدان أخرى.

وأخيرًا، فيما يتعلق بتطبيق الذكاء الاصطناعي في المجال العسكري، ستواصل اليابان المشاركة بنشاط وبشكل بنّاء في المناقشات الدولية بهدف تحقيق تفاهم مشترك في المجتمع الدولي من خلال مناقشات متوازنة تأخذ في الحسبان الاعتبارات الإنسانية والمنظورات الأمنية.

ليتوانيا

[الأصل: بالإنكليزية]
[9 نيسان/أبريل 2025]

تعرب ليتوانيا عن تقديرها للفرصة التي أتيحت لها لتقديم تقرير الأمين العام وفقا لقرار الجمعية العامة 239/79. وكان من دواعي سرور ليتوانيا أن تؤيد القرار الذي اتخذته الجمعية العامة في 24 كانون الأول/ديسمبر 2024.

وتشير ليتوانيا إلى أن تطوير الذكاء الاصطناعي واستخدامه في المجال العسكري يمثل فرصاً وتحديات للسلام والأمن الدوليين. وتعلق ليتوانيا أهمية كبيرة على وضع معايير ومبادئ للاستخدام المسؤول، من شأنها أن تسمح للدول بتسخير فوائد الذكاء الاصطناعي في المجال العسكري والتخفيف من حدة مخاطره المحتملة. وتؤمن ليتوانيا إيماناً راسخاً بأن جميع الدول المسؤولة لها مصلحة في ضمان التطبيق المسؤول للذكاء الاصطناعي في المجال العسكري. إن ليتوانيا مقتنعة بأن معالجة الآثار المترتبة على الذكاء الاصطناعي العسكري تتطلب العمل على الصعيد العالمي واتباع نهج متعدد أصحاب المصلحة، يشمل القطاعين العام والخاص والمجتمع المدني والأوساط الأكاديمية.

وتؤيد ليتوانيا بشدة الإعلان السياسي بشأن الاستخدام العسكري المسؤول للذكاء الاصطناعي وخاصية التشغيل الذاتي، الذي انضمت إليه ليتوانيا في 13 تشرين الثاني/نوفمبر 2023. ويتألف الإعلان السياسي من مبادئ غير ملزمة قانونًا وأفضل الممارسات لضمان الاستخدام المسؤول والقانوني للذكاء الاصطناعي في السياق العسكري. ويأخذ الإعلان السياسي في الاعتبار تدابير مثل استعراض القوانين، والرقابة المناسبة، وتقليل التحيز غير المقصود، وضمان أن تكون لقدرات الذكاء الاصطناعي العسكرية حالات استخدام واضحة ومحددة جيدًا. وتشجع ليتوانيا بقوة المزيد من الدول على التوقيع على الإعلان السياسي.

وعلاوة على ذلك، تشترك ليتوانيا في استراتيجية الذكاء الاصطناعي المتبعة في منظمة حلف شمال الأطلسي، والتي اعتُمدت في عام 2021 ونُقِّحت في عام 2024. وتحدد الاستراتيجية ستة مبادئ للاستخدام المسؤول للذكاء الاصطناعي في المجال العسكري، وهي: المشروعية، والمسؤولية والمساءلة،

وقابلية التفسير والتتبع، والموثوقية، وإمكانية التحكّم، والتقليل من التحيز. وتهدف هذه المبادئ غير الملزمة قانونًا، والتي تلتزم بها ليتوانيا، إلى تطبيقها على كامل الدورة العمرية لتطبيق الذكاء الاصطناعي.

وأخيرًا، يسر ليتوانيا أن توضح وجهات نظرها حول الفرص والتحديات التي يفرضها تطبيق الذكاء الاصطناعي في المجال العسكري على الأمن الدولي. وتعتقد ليتوانيا أن الذكاء الاصطناعي العسكري يمكن وينبغي أن يُستخدم على نحو مسؤول، أولاً وقبل كل شيء، لتعزيز الأمن القومي للدولة، وللمساهمة في تنفيذ القانون الدولي، بما في ذلك القانون الدولي الإنساني والوفاء بمختلف التزامات الدولة فيما يتعلق بحماية المدنيين. وبصرف النظر عن تعزيز حماية المدنيين في النزاعات المسلحة، يوفر الذكاء الاصطناعي المسؤول فرصًا لتحسين عملية صنع القرار واللوجستيات والتخطيط وغيرها من العمليات التي تؤدي إلى زيادة الكفاءة.

وفيما يتعلق بالمخاطر المحتملة للذكاء الاصطناعي في المجال العسكري، تود ليتوانيا تسليط الضوء على التحديات التي تشمل، على سبيل المثال لا الحصر، الأمن السيبراني، والتحيز غير المقصود في قدرات الذكاء الاصطناعي في المجال العسكري، والسلوك غير المقصود للمنظومات القائمة على الذكاء الاصطناعي. وتعتقد ليتوانيا أن أفضل طريقة للتصدي لهذه المخاطر المحتملة هي تنفيذ مبادئ الاستخدام المسؤول، بالإضافة إلى بناء القدرات والتدريب المناسب للموظفين على استخدام تطبيقات الذكاء الاصطناعي والمنظومات القائمة على الذكاء الاصطناعي. وتشدد ليتوانيا على أنه من أجل الحصول على فوائد الذكاء الاصطناعي في المجال العسكري واستخدام الذكاء الاصطناعي كقدرة دفاعية حاسمة، ينبغي للدول تجنب وضع قيود مفرطة لا داعي لها تعيق الابتكار في مجال الذكاء الاصطناعي، خاصة إذا رفضت الدول غير المسؤولة فرض أي قيود من هذا القبيل على الذكاء الاصطناعي العسكري.

المكسيك

[الأصل: بالإسبانية]
[10 نيسان/أبريل 2025]

الذكاء الاصطناعي ومنظومات الأسلحة الذاتية التشغيل والتحدي الذي يواجهه العالم فيما يتعلق بقواعدها التنظيمية

تقدم المكسيك هذه الوثيقة عملا بقرار الجمعية العامة 239/79، المعنون "الذكاء الاصطناعي في المجال العسكري وآثاره على السلام والأمن الدوليين".

وتقر المكسيك بأن تطبيق الذكاء الاصطناعي في المجال العسكري يمكن أن يكون له فوائد. ومع ذلك، فإنه يشكل أيضاً تحديات كبيرة للسلام والأمن الدوليين، الأمر الذي يتطلب اهتماماً عاجلاً ومنسقاً من المجتمع الدولي.

وتثمن المكسيك عمليات التبادل المتعددة الأطراف التي تجري في إطار الأمم المتحدة، مثل الحوار الافتتاحي بشأن الذكاء الاصطناعي العسكري والسلام والأمن، المعنون "الفرص والمخاطر والسلام والأمن الدوليان"، الذي ينسقه مكتب شؤون نزع السلاح، والذي يسهم في التوصل إلى فهم مشترك للمخاطر الناشئة والمسؤوليات المشتركة. وتتفق مع الرأي القائل بأن إدماج الذكاء الاصطناعي في المهام العسكرية يطرح

تحديات أساسية للسلام والأمن الدوليين، بما في ذلك التصعيد غير المقصود للنزاع، والغموض الاستراتيجي، وزيادة إمكانية التشغيل الذاتي في استخدام القوة.

وترى المكسيك أنه ينبغي إعطاء الأولوية لتعزيز التعاون الدولي وزيادة الشفافية وتبادل أفضل الممارسات وبناء القدرات التي تدعم ثقافة الامتثال التنظيمي واحترام القانون الدولي، فضلاً عن النهوض بعملية وضع أطر تنظيمية تضمن أن يكون تطوير الذكاء الاصطناعي ونشره في السياقات العسكرية محكوماً بمبادئ أخلاقية وقانونية وإنسانية، مما يحول دون أن تؤدي هذه التكنولوجيا إلى تعميق أوجه عدم التناظر أو تقويض الاستقرار الدولي.

السلام والأمن الدوليان

تعتقد المكسيك أن من الضروري اتخاذ إجراءات لمنع انتشار هذه التكنولوجيات وإساءة استخدامها، بما في ذلك من قبل جهات فاعلة من غير الدول وخارج اطر قانونية واضحة.

ويجب ألا يكون لإدماج التكنولوجيات الجديدة والناشئة في المجال العسكري الأسبقية على السلام والأمن الدوليين. ويجب أن يأخذ هذا الإدماج في الاعتبار التنمية البشرية والتمكين الاجتماعي، لا سيما لصالح البلدان النامية. ولذلك ينبغي أن تسترشد أهداف هذا النوع من التكنولوجيا بالاستخدامات السلمية وحل النزاعات، بدلاً من السعي وراء آلة عسكرية أكثر كفاءة.

إن التعقيد المتزايد للتهديدات الرقمية وإمكانية استخدام التقنيات الناشئة كوسيلة لشن هجمات من دولة إلى دولة، بالإضافة إلى صعوبة ضمان موثوقية ودقة المنظومات الذاتية التشغيل في السياقات العسكرية، والتعرض لأوجه الضعف طوال دورة حياة الذكاء الاصطناعي، وأشكال التحيز الخوارزمية، وتسمم البيانات، واستخدام النماذج التوليدية لأغراض خبيثة، تؤكد كلها الحاجة إلى اتخاذ إجراءات استباقية للتخفيف من للمخاطر.

والتقدم العلمي والتكنولوجي، خاصة في مجال الذكاء الاصطناعي والمنظومات الذاتية التشغيل والتكنولوجيات الرقمية والكمية، يتجاوز القدرة الحالية للأطر التنظيمية على إدارة مثل هذه المخاطر. ولذلك تكرر المكسيك الحاجة إلى وضع أطر حوكمة شاملة، وتعزيز التعاون الدولي والحوار المتعدد الأطراف، وإعطاء الأولوية للشفافية والمساءلة والتحكم البشري الفعلي طوال دورة حياة هذه التكنولوجيات، بما في ذلك الاختبارات الصارمة والضمانات الأخلاقية لنشرها.

وفي غياب الأطر القانونية الدولية الواضحة وتوافق الآراء الضروري بين الأطراف المتعددة، فإن استخدام مصطلح "مسؤول" في هذا السياق ينبغي ألا يفسر على أنه تأييد أو قبول ضمني لاستخدام أو تطوير القدرات العسكرية التي تدعم الذكاء الاصطناعي وتعمل بالتشغيل الذاتي. ومن الضروري أن يرتبط مبدأ المسؤولية بالشرعية والمساءلة.

وفي هذا الصدد، ترى المكسيك أن من الضروري إنشاء آليات حوكمة وآليات تنظيمية تقلل من احتمال استخدام الذكاء الاصطناعي وغيره من التكنولوجيات الكاسحة لأغراض عدائية، مع الاعتراف بأن المخاطر لا تنشأ فقط أثناء نشرها التشغيلي، بل منذ المراحل الأولى من التصميم والتطوير.

السياق التشغيلي

تشير المكسيك إلى أنه قد يكون للذكاء الاصطناعي آثار متباينة بالنظر إلى السياقات التشغيلية العسكرية المختلفة التي يمكن أن تُدمج فيها هذه التكنولوجيا.

وفي سياق النزاع المسلح، يجب التأكد من أن أي تكنولوجيا قائمة على الذكاء الاصطناعي تُستخدم وفقًا للقانون الدولي الإنساني، ولا سيما مبادئ التمييز والتناسب والحيطة والإنسانية.

وفي مجال عمليات حفظ السلام والاستجابة للكوارث، يمكن للذكاء الاصطناعي أن يقدم إسهامات إيجابية في التنسيق اللوجستي والتنبؤ بالمخاطر ورعاية السكان المتضررين، شريطة احترام إطار حقوق الإنسان احتراماً كاملاً.

وفيما يتعلق بأمن الحدود، تدرك المكسيك أن الذكاء الاصطناعي يمكن أن يعزز قدرات المراقبة؛ لكنها تشدد على أهمية ضمان احترام كرامة جميع الأشخاص وتجنب القرارات الآلية التي تديم الممارسات التمييزية.

منظومات الأسلحة الفتاكة الذاتية التشغيل

ترى المكسيك أن أحد الجوانب الرئيسية في هذه المناقشة هو منظومات الأسلحة الفتاكة الذاتية التشغيل، التي تشكل مصدر قلق خاص فيما يتعلق بالسلام والأمن الدوليين. وفي هذا الصدد، تشدد المكسيك على ضرورة عدم تجزئة المناقشات المتعددة الأطراف بشأن إدماج التكنولوجيات الجديدة في المجال العسكري، وتعتقد أن منظومات الأسلحة الفتاكة الذاتية التشغيل يجب أن تكون جزءاً لا يتجزأ من هذه المناقشات.

وترى المكسيك أنه من الملح أن يوضح المجتمع الدولي المحظورات والقواعد التنظيمية بشأن منظومات الأسلحة الفتاكة الذاتية التشغيل، نظراً لتعارضها مع القانون الدولي الإنساني ومخاطرها الأخلاقية والقانونية والأمنية.

وقد دعمت المكسيك القرارين 241/78 و 62/79 بشأن منظومات الأسلحة الفتاكة الذاتية التشغيل وشاركت في تقديمهما في الجمعية العامة، بهدف إيجاد مساحة مشروعة متعددة الأطراف للتصدي لهذه التحديات.

وتؤيد المكسيك دعوة الأمين العام واللجنة الدولية للصليب الأحمر إلى الشروع في مفاوضات بشأن صك ملزم قانونا ينشئ المحظورات والقواعد التنظيمية اللازمة بشأن منظومات الأسلحة الفتاكة الذاتية التشغيل بحلول عام 2026، على النحو المنصوص عليه في الخطة الجديدة للسلام.

وقد أبدت المكسيك التزامها السياسي بهذه القضية من خلال مشاركتها في مؤتمر سان خوسيه (2023)، وتمسكها ببيان بيليم ومشاركتها النشطة في المؤتمر المعنون "البشرية أمام مفترق طرق: منظومات الأسلحة الذاتية التشغيل والتحديات المتصلة بتنظيمها" (فيينا، 2024)، الذي أقرت المكسيك تقريره الختامي.

وتعتقد المكسيك أن منظومات الأسلحة الفتاكة الذاتية التشغيل تشكل عددًا من المخاطر، بما في ذلك أنها:

- تستبعد الحكم البشري من القرارات الحاسمة بشأن استخدام القوة.

- تحل محل التقييم السياقي الذي لا غنى عنه في العمليات العسكرية.
- تضعف آليات المساءلة وإسناد المسؤولية.

ولا ينبغي أبدا نقل المسؤولية عن استخدام القوة إلى الآلات. ويجب أن تبقى القرارات المتعلقة بنشر المنظومات المسلحة أو تفعيلها أو تجاوزها دائماً بيد أشخاص بشر يخضعون للمسؤولية القانونية.

وتكرر المكسيك التأكيد على أن جميع أشكال التكنولوجيا العسكرية، بما فيها تلك القائمة على الذكاء الاصطناعي، يجب أن تحترم الالتزامات الدولية المنبثقة عن:

- ميثاق الأمم المتحدة
- القانون الدولي الإنساني
- القانون الدولي لحقوق الإنسان
- القانون الجنائي الدولي
- قانون المسؤولية الدولية

وفي هذا الصدد، ترى المكسيك أنه من الأهمية بمكان حظر منظومات الأسلحة التي تستخدم فيها تكنولوجيا:

- لا يمكنها التمييز بين الأهداف العسكرية والمدنية
- لا يمكنها تطبيق مبدأ التناسب على الأضرار الجانبية
- ليس فيها آليات إلغاء إذا تبين أن الهجوم غير مبرر
- تسبب معاناة لا مبرر لها أو إصابة مفرطة

وتصر المكسيك على الحاجة الملحة للبدء في مفاوضات بشأن صك ملزم قانوناً يحدد المحظورات والقواعد التنظيمية بشأن منظومات الأسلحة الفتاكة الذاتية التشغيل؛ ويضمن الحفاظ على التحكم البشري الفعلي بالأنشطة الأساسية؛ ويشمل آليات فعالة للتنفيذ والرصد والمساءلة.

الفوائد والمخاطر

فيما يتعلق بالاستخدامات المحددة للذكاء الاصطناعي، تدرك المكسيك كلا من الفوائد والمخاطر في المجالات التالية:

- القيادة والتحكم: في ظل ظروف معينة، قد يحسّن الذكاء الاصطناعي من كفاءة قرارات العمليات، ولكن يجب أن تظل هذه القرارات خاضعة لتحكم بشري كبير، خاصة عندما تتعلق باستخدام القوة. ويتسم الذكاء الاصطناعي بالقدرة على معالجة وتحليل كميات كبيرة من البيانات والمعلومات، بما يتجاوز القدرات البشرية بكثير، ويجعل من الممكن تسريع التنبؤ بالاتجاهات المستقبلية وتسهيله وتبسيطه وتوجيه القرارات الاستراتيجية في الوقت الفعلي.

- العمليات السيبرانية: يوفر الذكاء الاصطناعي قدرات قيّمة للتنبؤ بالحوادث السيبرانية والاستجابة لها، ولكنه يزيد أيضاً من مخاطر تصاعد التوترات، بما في ذلك استخدام التشغيل الذاتي في الهجوم دون رقابة كافية.

- إدارة المعلومات والخدمات اللوجستية: يمكن أن تيسر معالجة البيانات الضخمة من خلال الذكاء الاصطناعي اتخاذ القرارات في الوقت الفعلي، ولكن يجب أن يتم ذلك بما يتماشى مع بروتوكولات تضمن الاستخدام الأخلاقي والمبرر والمسؤول للذكاء الاصطناعي.

وبغض النظر عما سبق، تؤكد المكسيك على المخاطر التكنولوجية المرتبطة بإدماج الذكاء الاصطناعي في السياقات العسكرية، نظرًا لأن الأدلة تشير إلى استمرار وجود أعطال تقنية أو أخطاء غير متوقعة يمكن أن تؤدي إلى تصعيد النزاع.

هولندا (مملكة -)

[الأصل: بالإنكليزية]
[7 نيسان/أبريل 2025]

ترحب مملكة هولندا بفرصة تقديم آرائها، وفقاً للقرار 239/79، الذي اتخذته الجمعية العامة في 24 كانون الأول/ديسمبر 2024، بشأن التحديات والفرص التي يطرحها الذكاء الاصطناعي في المجال العسكري على السلم والأمن الدوليين.

وتدرك هولندا التطبيقات العسكرية المحتملة للذكاء الاصطناعي وتلتزم بتطوير الذكاء الاصطناعي ونشره واستخدامه في المجال العسكري بشكل مسؤول. ويتمثل الموقف الأساسي لهولندا في أن تطبيق الذكاء الاصطناعي في المجال العسكري يجب أن يكون متوافقاً مع القانون الدولي، بما في ذلك ميثاق الأمم المتحدة والقانون الدولي الإنساني والقانون الدولي لحقوق الإنسان.

وفي يومي 15 و 16 شباط/فبراير 2023، استضافت هولندا أول قمة لتسخير الذكاء الاصطناعي المسؤول في المجال العسكري. ومنذ ذلك الحين، وفرت عملية الذكاء الاصطناعي المسؤول في المجال العسكري منتدى من منتديات أصحاب المصلحة المتعددين لممثلي الحكومات والمؤسسات المعرفية ومراكز الأبحاث وقطاع الصناعة ومنظمات المجتمع المدني لمناقشة الفرص والتحديات الرئيسية المرتبطة بالتطبيقات العسكرية للذكاء الاصطناعي. وتجري المناقشات سنوياً على المستوى العالمي وعلى مدار السنة خلال الفعاليات الإقليمية المعنية بالذكاء الاصطناعي المسؤول في المجال العسكري التي استضافتها حتى الآن سنغافورة وكينيا وتركيا وشيلي وهولندا.

وفي قمة عام 2023، اتفقت هولندا و 57 من البلدان الأخرى على توجيه نداء مشترك للعمل على تطوير ونشر واستخدام الذكاء الاصطناعي في المجال العسكري على نحو مسؤول. وفي عام 2024، أيدت هولندا مخطط العمل الذي تم الاتفاق عليه خلال قمة تسخير الذكاء الاصطناعي المسؤول في المجال العسكري لعام 2024، التي استضافتها جمهورية كوريا وشاركت هولندا في استضافتها. وبالإضافة إلى ذلك، صادقت هولندا على الإعلان السياسي بشأن الاستخدام العسكري المسؤول للذكاء الاصطناعي وخاصية التشغيل الذاتي.

وخلال قمة عام 2023، أطلقت هولندا الهيئة العالمية لتسخير الذكاء الاصطناعي المسؤول في المجال العسكري وكلفتها بوضع توصيات قصيرة وطويلة الأجل للحكومات ومجتمع أصحاب المصلحة المتعددين على نطاق أوسع. وتنتظر هولندا نشر تقرير عن التوجيهات الاستراتيجية لهذه الهيئة في أيلول/سبتمبر 2025.

ويلخص الفرع الوارد أدناه الموقف الهولندي ويحدد المسائل الرئيسية التي تتطلب زيادة النظر فيها.

الفرص المتاحة لتحقيق السلام والأمن الدوليين

من منظور عسكري، تتمثل الفوائد الأساسية للذكاء الاصطناعي في السرعة والنطاق. وتتيح تكنولوجيا الذكاء الاصطناعي معالجة البيانات وتحليلها بوتيرة أسرع بكثير. كما تساعد أنظمة وضع السيناريوهات ودعم اتخاذ القرارات القائمة على الذكاء الاصطناعي القادة على صياغة مسارات العمل. ويؤدي ذلك إلى تحسين الرؤية الاستراتيجية والقدرة على الاستجابة للتهديدات بسرعة وفعالية.

وتعتقد هولندا أن الذكاء الاصطناعي يمكن أن يساهم أيضًا في تحقيق السلام والأمن الدوليين من خلال تقديم رؤية أعمق، وتحسين الربط، وتعزيز حماية المدنيين، والحد من المخاطر أثناء عمليات الخطوط الأمامية:

- تعمل أنظمة التحليل ودعم اتخاذ القرارات القائمة على الذكاء الاصطناعي على تحسين الحالة المعلوماتية لدى القادة، فيما يتعلق بكل من الوضع على الأرض والتطورات الاستراتيجية الطويلة المدى. ويساعد ذلك في تحسين الرؤية حول ديناميات السكان المدنيين داخل مناطق النزاع، والتحديات الأمنية المناخية، والعنف الجنساني، والأنماط السلوكية للمنظمات الإرهابية. ويمكن استخدام هذه المعلومات بدورها لتحسين إدارة المخاطر والنزاعات، وبالتالي المساهمة في تحقيق السلام والأمن الدوليين.

- ترى هولندا أن من المفيد تطبيق الذكاء الاصطناعي في المجال العسكري لتحسين الربط بين قوات الدفاع، وبين قوات الدفاع والجهات الفاعلة الأخرى من قبيل الجهات الفاعلة في مجال تقديم المساعدة الإنسانية ومنظمات الرصد والحكومات المحلية. ويمكن تبادل البيانات بين عدد كبير من المستخدمين، مما يؤدي إلى إنشاء "مصادر وحيدة للحقيقة" باستخدام أجهزة استشعار "ذكية" تعمل في بيئة آمنة ومترابطة شبكياً. ويمكن أيضًا الاستعانة بعملاء الذكاء الاصطناعي لتبادل البيانات بسرعات عالية بشكل متزايد. ويعود تحسين الربط من خلال تحسين تبادل البيانات بسرعة أعلى بالفائدة على السلام والأمن الدوليين من خلال تعزيز الاتصال وتبادل المعلومات والتعاون الدولي، على سبيل المثال في نظم الإنذار المبكر وإدارة الأزمات.

- تولي هولندا أهمية كبيرة لإمكانات الذكاء الاصطناعي في مجال حماية المدنيين. ويمكن للذكاء الاصطناعي التعرف على الأنماط والانحرافات في كميات كبيرة من البيانات، مما يتيح فهماً أشمل للبيئة المدنية. ويمكن لهذا الفهم المتزايد أن يقلل من خطر وقوع الخطأ في تحديد الهوية والأضرار الجانبية والخسائر في صفوف المدنيين. وعلى نطاقٍ أوسع، يوفر الذكاء الاصطناعي إمكانية تحسين عملية التعرف على التهديدات المحتملة للمدنيين والأعيان المدنية، مما يمكّن القوات المسلحة من الاستجابة بسرعة وبشكل مناسب. كما يمكن للذكاء الاصطناعي المساعدة في تحسين

جهود المساعدة الإنسانية مثل توفير الغذاء والمأوى والرعاية الطبية في مناطق النزاع. وأخيرًا، يمكن للذكاء الاصطناعي تحسين التحقيقات في الخسائر في صفوف المدنيين من خلال جمع البيانات والأدلة وتحليلها، من أجل تحديد سبب الضرر وضمان إمكانية محاسبة المسؤولين عنه.

- يقلل الذكاء الاصطناعي من المخاطر التي يتعرض لها الأفراد العسكريون في الخطوط الأمامية، لأن المنظومات الذاتية التشغيل القائمة على الذكاء الاصطناعي قد تحل محل البشر في بعض الأنشطة في المناطق الصعبة أو الخطرة. ومن الأمثلة على ذلك المراقبة تحت الماء ودعم عمليات البحث والإنقاذ في ظل الأحوال الجوية القاسية. وقد يساعد الذكاء الاصطناعي أيضًا في تقليل التكاليف الطبية وتكاليف إعادة التأهيل من خلال تقليل تعرض الأفراد العسكريين لبيئات عالية الخطورة.

التحديات التي تواجه السلام والأمن الدوليين

تحدد هولندا مختلف المخاطر التي تهدد السلم والأمن الدوليين والتي تنشأ عن تطبيق الذكاء الاصطناعي في المجال العسكري:

- تشعر هولندا بالقلق من إمكانية استخدام الذكاء الاصطناعي لتضخيم الهجمات السيبرانية والتلاعب بالمعلومات وتحسينها وأتمتتها، وكلاهما يقوض السلام والأمن الدوليين. ومع ظهور الذكاء الاصطناعي التوليدي، أصبح تنفيذ التلاعب بالمعلومات والهجمات السيبرانية الآلية أسهل. وعند نشرها في المجال العسكري، فإنها تعطل خطوط الاتصال العملياتية وتعقّد عملية اتخاذ القرار. وعلى المدى الطويل، يمكن أن يؤدي نشر المعلومات المضللة والهجمات السيبرانية الآلية على نطاق واسع إلى تآكل الثقة في خطوط الاتصال العسكرية. كما أنها يمكن أن تؤثر على الثقة بين الدول، وبالتالي يمكن أن تضر بالعلاقات الهشة، لا سيما بين الدول التي توجد بالفعل على شفا نزاع محتمل.

- يمكن أن تؤدي المخاطر المرتبطة بتطبيق الذكاء الاصطناعي في المجال العسكري إلى منظومات قد تنتهك القانون الدولي. وقد تحدث أوجه القصور هذه بسبب عدم كفاية التكيف مع السياق والبيانات والمصطلحات العسكرية، وتؤدي بدورها إلى تبسيط مفرط في عملية اتخاذ القرار العسكري أو تجاهل سياقات عملياتية محددة، على سبيل المثال. ومن المحتمل أيضًا أن تنتهك الدول الالتزامات القانونية الدولية إذا كان التطبيق يتصرف بشكل غير متوقع، أو تترتب عليه نتائج تمييزية استنادًا إلى خصائص غير ذات صلة، أو يقترح مسارات عمل غير قانونية. ونظرًا للانتشار المتزايد للذكاء الاصطناعي، فإن تأثير تحيز الآلة والتحيز في مجموعات البيانات والقرارات البشرية القائمة على منظومات الذكاء الاصطناعي غير الملائمة يمكن أن يخلق تحديات كبيرة في تحديد المسؤولية وضمان المساءلة والمعالجة المناسبة. والأهم من ذلك، لا يمكن أن نتوقع من تطبيقات الذكاء الاصطناعي أن تفكر أو تعمل بنفس الطريقة التي يعمل بها البشر.

- يشكل خطر التصعيد القائم على الذكاء الاصطناعي مخاطر محتملة على السلم والأمن الدوليين. ومع تسريع الذكاء الاصطناعي لحلقة "الملاحظة والتوجيه والقرار والفعل" من خلال زيادة السرعة وقدرات النطاق، قد تنشأ تصورات خاطئة بسبب التفاوتات بين النوايا العسكرية والتحليلات التي تنتجها المنظومات القائمة على الذكاء الاصطناعي. ولذلك، يمكن أن يساهم الذكاء الاصطناعي

عن غير قصد في التصعيد. ونظراً لأن منظومات الذكاء الاصطناعي قادرة على تحديد الأهداف المحتملة بسرعة ونطاق يفوقان قدرة البشر، فإن استخدامها قد يزيد أيضاً من حدة النزاعات وقوة فتكها.

- نتيجة لذلك، يمثل إنشاء منظومات دفاعية قوية تحدياً متزايد الأهمية. إن السرعة التي تظهر بها تطبيقات الذكاء الاصطناعي الجديدة تجعل من الصعب تنفيذ استراتيجيات وتكتيكات لمواجهتها والتصدي لها بفعالية في السياق العسكري. وهذه النتيجة المحددة للاستخدام المتزايد لمنظومات الذكاء الاصطناعي يمكن أن تكون في صالح الأعمال الهجومية، وبالتالي تؤثر سلبًا على السلم والأمن الدوليين.

- مع حصول المنظمات الإرهابية وشبكات الجريمة المنظمة وغيرها من الجهات من غير الدول على قدرات الذكاء الاصطناعي العسكرية، فإن زعزعة الاستقرار تشكل مصدر قلق آخر. وفي هذا السياق، تشعر هولندا بالقلق من أن الذكاء الاصطناعي يمكن أن يجعل إنتاج الأسلحة الكيميائية والبيولوجية والإشعاعية والنووية في متناول هذه الجهات.

ونظراً للتطور السريع لتكنولوجيات الذكاء الاصطناعي، تقر هولندا بأنه لا يمكن التنبؤ بالتحديات والفرص المتعلقة بالسلام والأمن الدوليين بشكل كامل في الوقت الحاضر. فبعضها جديد تمامًا، والبعض الآخر موجود بالفعل ولكن قد يتفاقم بسبب تطبيق الذكاء الاصطناعي. إن الحوار الدولي المستمر حول هذه المسألة ضروري لضمان التطبيق المسؤول للذكاء الاصطناعي في المجال العسكري من قبل جميع الدول.

التطبيق المسؤول للذكاء الاصطناعي في المجال العسكري

من أجل ضمان تطبيق الذكاء الاصطناعي بشكل مسؤول في المجال العسكري، يجب الإبقاء على التقدير والتحكم البشريين بما يناسب السياق. ويجب أن يظل البشر مسؤولين وخاضعين للمساءلة. ومع ذلك، من المهم ملاحظة النقاط الموضحة أدناه.

المزيد من التحكم البشري لا يضمن المزيد من الذكاء الاصطناعي المسؤول

تعتقد هولندا أنه لا توجد طريقة واحدة تناسب الجميع لدمج ما يكفي من التقدير والتحكم البشريين في تطبيقات الذكاء الاصطناعي. ويتراوح التقدير والتحكم البشريان من التحكم البشري المباشر إلى مستويات أعلى من الأتمتة والتشغيل الذاتي، وفقا لعدة عوامل. لذلك يجب تحديد الدرجة المطلوبة من التقدير والتحكم البشريين التي يجب أن يمارسها البشر على التطبيقات والمنظومات القائمة على الذكاء الاصطناعي على أساس كل حالة على حدة. وهذه هي الطريقة الوحيدة لمراعاة عوامل متعددة مثل السياق التشغيلي، والتأثير على قدرة التكنولوجيا على العمل بشكل مستقل في البيئات المعادية، وبارامترات النظام والتفاعل بين الإنسان والآلة.

البحث والتطوير أمر أساسي للنشر المسؤول لتطبيقات الذكاء الاصطناعي في المجال العسكري

تؤمن هولندا بأهمية البحث والتطوير. ويجب على الدول أن تقيّم بشكل كافٍ ما إذا كانت تطبيقات الذكاء الاصطناعي تعمل بالطريقة التي صُممت من أجلها ويمكن نشرها في سياق استخدام محدد. وهذا ضروري بشكل خاص أثناء القتال وفي البيئات الأخرى العالية المخاطر. ومن خلال البحث والتطوير

بالمعنى العام، ومن خلال إجراءات الاختبار والتقييم والتحقق والمصـــادقة المثبتة والموثوقة فيما يخص تطبيقات محددة للذكاء الاصـطناعي، يمكن اكتشـاف المشـكلات المحتملة والتخلص منها أو التخفيف من حدتها قبل النشر. وبالإضافة إلى ذلك، من المهم أن يتم تدريب الأفراد العسكريين بشكل كافٍ على تطبيقات الذكاء الاصـطناعي وتعريفهم بها قبل نشـر التطبيقات، لضـمان فهمهم لقدرات التطبيقات وأوجه قصـورها. ويكتسب هذا الأمر أهمية خاصة نظراً للتطورات التكنولوجية السريعة المتعلقة بتطبيقات الذكاء الاصطناعي وحقيقة أن استخدامها أصبح أقل تكلفة.

يجب أن تكون الحوكمة الدولية للذكاء الاصطناعي العسكري مرنة وشاملة وواقعية

فيما يتعلق بالحوكمة الدولية للذكاء الاصـطناعي في المجال العسـكري، تدرك هولندا الحاجة إلى اتباع نهج مرن ومتوازن وواقعي. أولاً، يجب أن تتســم الأطر المنظِّمة بالمرونة من أجل مواكبة التطورات الســريعة فيما يخص التكنولوجيا وميادين القتال. ثانياً، تحتاج الأطراف إلى العمل على التوصـــل إلى فهم مشـترك للذكاء الاصـطناعي في المجال العسـكري والفرص والمخاطر والحلول المحتملة التي تصـاحبه. وسـيتطلب ذلك حوارًا عالميًا شـاملاً ومشـاركة فعالة من جميع مجموعات أصـحاب المصـلحة، بما في ذلك الدول والمؤسـسـات المعرفية والمجتمع المدني وقطاع الصـناعة. ثالثًا، ينبغي للدول التركيز على وضـع ضـمانات للتطبيق المسؤول للذكاء الاصـطناعي في المجال العسكري، على سبيل المثال، مع التركيز على قضـــايا مثل ضـــمان قابلية التتبع أو قابلية الفهم. رابعًا، يجب أن تراعي الحوكمة الدولية لنشـــر الذكاء الاصـطناعي العسكري وجهات نظر الدول المختلفة بشأن القواعد التنظيمية. وفي إطار بارامترات الالتزامات القانونية القائمة، ينبغي ألا تعيق الحوكمة الدولية للذكاء الاصطناعي في المجال العسكري قدرات الدول على الابتكار.

مناقشة حول منظومات الأسلحة الذاتية التشغيل

بما أن للذكاء الاصـطناعي إمكانات كبيرة لتشغيل منظومات الأسلحة الذاتية التشـغيل، فهناك أوجه تشـابه واضـحة بين النقاش الأوسـع حول اسـتخدامه في المجال العسـكري والنقاش حول تنظيم منظومات الأسلحة الذاتية التشغيل. وترى هولندا أن المناقشات الدولية حول هذين الموضوعين متكاملة ويفيد بعضها بعضا.

نيوزيلندا

[الأصل: بالإنكليزية]
[11 نيسان/أبريل 2025]

يرد هذا التقرير الوطني المقدم من نيوزيلندا على المذكرة الشفوية المؤرخة 12 شباط/فبراير 2025 من مكتب شؤون نزع السلاح، وينبغي قراءته جنباً إلى جنب مع رد نيوزيلندا على المذكرة الشفوية للمكتب المؤرخة 1 شباط/فبراير 2024[1].

(1) متاح على الرابط التالي: www.mfat.govt.nz/assets/Peace-Rights-and-Security/Disarmament/New-Zealand-submission-to-the-UN-Secretary-General-on-autonomous-weapon- systems.pdf.

موقف نيوزيلندا من الذكاء الاصطناعي في المجال العسكري

تدرك نيوزيلندا أن التطبيقات المحتملة والحالية للذكاء الاصطناعي في المجال العسكري سيكون لها آثار بعيدة المدى ومتعددة الأوجه.

وفي حين أنه من غير الواضح حتى الآن طبيعة ومدى العديد من هذه الآثار، إلا أن الذكاء الاصطناعي يُطبق بالفعل في مجموعة واسعة من الوظائف العسكرية من قبل بعض المنظمات العسكرية، بما في ذلك الاستخبارات والتخطيط واللوجستيات والملاحة والاتصالات. وعلى الرغم من أن الذكاء الاصطناعي ينطوي على بعض المخاطر، إلا أنه يمكن أن يمنح المستخدمين مزايا كبيرة في المجال العسكري، بما في ذلك زيادة السرعة والكفاءة والدقة والإلمام بالحالة. وكغيرها من الجيوش الأخرى، تعتزم قوات الدفاع النيوزيلندية اغتنام الفرص التي يتيحها الذكاء الاصطناعي لتحسين عملياتها والحفاظ على قابلية التشغيل البيني مع شركائها.

ونعيد تأكيد الفقرة 1 من قرار الجمعية العامة 239/79، أي "أن القانون الدولي، بما في ذلك ميثاق الأمم المتحدة والقانون الدولي الإنساني والقانون الدولي لحقوق الإنسان، ينطبق على ما يخضع لأحكامه من مسائل تطرأ في جميع مراحل الدورة العُمرية للذكاء الاصطناعي، بما في ذلك المنظومات التي تستمد فعاليتها من الذكاء الاصطناعي، في المجال العسكري". وبالإضافة إلى الالتزامات القانونية الملزمة، ينبغي مراعاة المعايير الأخلاقية ذات الصلة طوال الدورة العمرية للذكاء الاصطناعي في المجال العسكري.

وتعترف نيوزيلندا بأن الذكاء الاصطناعي وثيق الصلة بتطوير واستخدام بعض منظومات الأسلحة، وذلك في مجالات منها على سبيل المثال الارتقاء بمستويات التشغيل الذاتي. ويرد موقف نيوزيلندا من منظومات الأسلحة الذاتية التشغيل بالتفصيل في ردها على المذكرة الشفوية لمكتب شؤون نزع السلاح المؤرخة 1 شباط/فبراير 2024.

ومن الوارد أن يتم تطبيق الذكاء الاصطناعي في مجال تطوير أسلحة الدمار الشامل. ومن الواضح أن الأسلحة البيولوجية والكيميائية محظورة بموجب القانون الدولي، وتؤكد نيوزيلندا أن معيار الغرض العام في كل من اتفاقية حظر استحداث وإنتاج وتخزين الأسلحة البكتريولوجية (البيولوجية) والتكسينية وتدمير تلك الأسلحة واتفاقية حظر استحداث وإنتاج وتخزين واستعمال الأسلحة الكيميائية وتدمير تلك الأسلحة ينطبق في حالة استخدام الذكاء الاصطناعي لتطوير مثل هذه الأسلحة، مما يعني، في جملة أمور، أن الذكاء الاصطناعي يجب ألا يستخدم لهذا الغرض. وبالإضافة إلى ذلك، وكما أشارت إلى ذلك الدول الأطراف في معاهدة حظر الأسلحة النووية، بما فيها نيوزيلندا، من الضروري الإبقاء على التحكم البشري الفعلي في الأسلحة النووية ومنظومات إيصالها، ريثما تتم إزالتها وتحقيق عالم خال من الأسلحة النووية.

المقترحات المعيارية الحالية والناشئة

يعد التوصل إلى فهم مشترك وبناء المعايير من الجوانب المهمة لتعزيز الاستخدام العسكري المسؤول للذكاء الاصطناعي. وفي عام 2024، انضمت نيوزيلندا إلى الإعلان السياسي بشأن الاستخدام المسؤول للذكاء الاصطناعي والتشغيل الذاتي للأغراض العسكرية بقيادة الولايات المتحدة إلى جانب العديد من البلدان الأخرى. ويؤكد الإعلان على أن "الاستخدام العسكري للذكاء الاصطناعي يمكن، بل وينبغي، أن

يكون أخلاقيا ومسؤولا وأن يعزز الأمن الدولي". وشاركت نيوزيلندا أيضاً في مؤتمرات قمة تسخير الذكاء الاصطناعي المسؤول في المجال العسكري.

وترى نيوزيلندا أن من المفيد إجراء مناقشات متعددة الأطراف، بما في ذلك من خلال الأمم المتحدة، مكرسة لوضع معايير حول الذكاء الاصطناعي في المجال العسكري والاتفاق عليها. ومن المهم مشاركة أصحاب المصلحة من غير الدول، بما في ذلك المجتمع المدني والمنظمات الدولية والإقليمية وقطاع الصناعة في هذه المناقشات في جميع مراحل هذه العمليات.

النرويج

[الأصل: بالإنكليزية]
[11 نيسان/أبريل 2025]

ترحب النرويج بالفرصة التي أتيحت لها لتقديم آرائها بشأن الفرص والتحديات التي يطرحها تطبيق الذكاء الاصطناعي في المجال العسكري على السلام والأمن الدوليين، مع التركيز بشكل خاص على مجالات أخرى غير منظومات الأسلحة الفتاكة الذاتية التشغيل، عملا بقرار الجمعية العامة 239/79 المعنون "الذكاء الاصطناعي في المجال العسكري وآثاره على السلام والأمن الدوليين".

وكما هو معترف به في الموجز السياساتي للأمين العام الصادر في تموز/يوليه 2023 حول خطة جديدة للسلام، فإن الذكاء الاصطناعي هو تكنولوجيا تمكينية وكاسحة في آن واحد يتم استخدامها بشكل متزايد في طائفة واسعة من التطبيقات المدنية والعسكرية وذات الاستخدام المزدوج. ويشكل شيوع استعمال الذكاء الاصطناعي بشكل متزايد في كل المجالات، بالاقتران مع قابلية استخدامه الموسع بسرعة وافتقاره للشفافية ووتيرة الابتكار المتصلة به، مخاطر محتملة على السلام والأمن الدوليين، ويطرح تحديات في مجال الحوكمة.

وباعتبار النرويج من المدافعين الدائمين عن القانون الدولي والتعددية والابتكار المسؤول في قطاع الدفاع، فإنها تدعم الجهود الرامية إلى تشجيع التوصل إلى فهم مشترك وتعزيز الحوكمة ووضع قواعد تنظيمية مناسبة للذكاء الاصطناعي في المجال العسكري. وكنقطة انطلاق، يجب تطوير تطبيقات الذكاء الاصطناعي في المجال العسكري ونشرها وتطبيقها بطريقة مسؤولة طوال دورتها العمرية الكاملة، وبما يتوافق مع القانون الدولي الساري، ولا سيما القانون الدولي الإنساني.

والأهم من ذلك، أكدت الجمعية العامة في قرارها 239/79 على انطباق القانون الدولي، بما في ذلك ميثاق الأمم المتحدة والقانون الدولي الإنساني والقانون الدولي لحقوق الإنسان، في مجال استخدام الذكاء الاصطناعي في المجال العسكري، وشددت على أهمية استخدام الذكاء الاصطناعي بشكل مسؤول يتمحور حول الإنسان.

وينطوي الذكاء الاصطناعي باعتباره تكنولوجيا تمكينية على إمكانات غير عادية لتحويل كل جانب من جوانب الشؤون العسكرية، بما في ذلك المشتريات والأجهزة والبرمجيات والعمليات والقيادة والتحكم والاتصالات الاستراتيجية والمراقبة والاستخبارات والتدريب وإدارة المعلومات والدعم اللوجستي. ويتيح تطبيق الذكاء الاصطناعي في المجال العسكري فرصا ويشكل مخاطر بشكل متوقع وغير متوقع في آن واحد على المستويين التكتيكي والاستراتيجي. ويمثل الذكاء الاصطناعي بوصفه تكنولوجيا ذات أغراض عامة عاملا

مضاعفا للقوة ذا قدرة على إعادة تشكيل إدارة الحرب. ويضيف التقارب التكنولوجي بين الذكاء الاصطناعي والتكنولوجيا العصبية والبيولوجيا التركيبية والحوسبة الكمية المزيد من التعقيد.

ولا بد أن يتم تطوير الذكاء الاصطناعي ونشره واستخدامه وإدارته بشكل مسؤول، بما يتماشى مع المبادئ الأخلاقية الأساسية، في ظل الامتثال الصارم لالتزامات الدول بموجب القانون الدولي، بما في ذلك القانون الدولي الإنساني والقانون الدولي لحقوق الإنسان، ويقع في صميم ذلك تحديد المخاطر والتخفيف من حدتها.

وتحدد الاستراتيجية النرويجية للذكاء الاصطناعي في قطاع الدفاع (2023) المجالات الرئيسية التي يمكن أن يساهم فيها الذكاء الاصطناعي بشكل بنّاء في مجالات أخرى غير منظومات الأسلحة الفتاكة الذاتية التشغيل:

- **تعزيز الإلمام بالحالة ودعم القرارات** – إن استخدام الذكاء الاصطناعي وارد وضروري في آن واحد في مجال الاستخبارات والمراقبة والاستطلاع، حيث لا يمكن تحليل كميات كبيرة ومتزايدة من البيانات يدوياً. ويمكن استخدام الذكاء الاصطناعي لفرز البيانات ذات الصلة، بسبل منها على سبيل المثال المعالجة المسبقة للبيانات، والترجمة الآلية أو الكشف عن أشياء خاصة في الصور، والكشف عن الحالات الشاذة والتكرار، بالإضافة إلى التحقق من المعلومات من مصادر مختلفة للكشف عن محاولات التضليل. ويمكن أن تؤدي التحسينات في هذا المجال إلى عمليات أكثر فعالية ودقة وإلى تقليل الخسائر في الأرواح.

- **الدفاع السيبراني** – تؤدي الرقمنة والاعتماد المتزايد على تكنولوجيا المعلومات والاتصالات إلى ظهور نقاط ضعف إلى جانب فوائدها. ويتيح الفضاء الرقمي للجهات الفاعلة التي تشكِّل تهديداً فرصة لاختراق البيانات، والانخراط في أعمال التجسس والتخريب، والقيام بحملات لتعزيز النفوذ. ويمكن للذكاء الاصطناعي أن يدعم قدرة قطاع الدفاع على اكتشاف التهديدات الرقمية ومراقبتها والإبلاغ عنها وإدارتها ومواجهتها. ويمكن أن يؤدي استخدام الذكاء الاصطناعي إلى جملة أمور منها تقديم صورة أشمل للأهداف والعلاقات المعقدة بسرعة أكبر، وجمع المعلومات من المصادر ذات الصلة، وتبسيط استخدام التحليل. ولا بد من تطوير المعرفة والخبرة فيما يتعلق بكيفية تشكيل الذكاء الاصطناعي تهديداً رقمياً لكي يتسنى كشف الهجمات الرقمية وتجنبها في المستقبل. ولذلك يجب أن يكون الذكاء الاصطناعي عنصراً محورياً في مواصلة تطوير عملية تصدي القطاع للتهديدات الرقمية، سواء باستخدام الأدوات الحالية أو المستقبلية.

- **اللوجستيات** – تعتمد العمليات العسكرية الناجحة والفعالة على الدعم اللوجستي الفعال. ومن خلال تبسيط الخدمات اللوجستية باستخدام المنظومات التي تعتمد الذكاء الاصطناعي، يمكن ضمان قدرة تشغيلية أفضل وتأهب أكبر. وقد قطعت تطبيقات الذكاء الاصطناعي في قطاع الخدمات اللوجستية المدنية شوطاً كبيراً بالفعل. ومن المحتمل أن يتم تكييف العديد منها بسهولة لاستخدامها في القطاع العسكري.

- **أنشطة الدعم** – من المحتمل أن يتم تحسين العديد من أنشطة الدعم العسكري وتبسيطها باستخدام الذكاء الاصطناعي. ويشمل ذلك المهام التي تدعم وتعزز القدرة التشغيلية، مثل تشغيل وصيانة العتاد وشراء العتاد والمباني وإدارتها والتخلص منها، وتعيين الموظفين وتدريبهم وإدارة شؤونهم،

وتقديم الخدمات المشتركة، مثل المحاسبة والأرشفة. وللذكاء الاصطناعي القدرة على تعزيز أنشطة الدعم من خلال تحسين استخدام البيانات للتحليلات ودعم اتخاذ القرارات، وأتمتة المهام، وتحسين القدرة على تناول المعلومات والمعارف. وهذا يمكن أن يجعل من الممكن التحول إلى نموذج الصيانة التنبؤية، وتحسين تدفق المعلومات، واستحداث أنظمة دعم جديدة وأفضل لإدارة الموارد البشرية، وتحسين نمذجة اتجاهات التكلفة للعتاد والمباني. وبالتالي، يمكن أن يؤدي استحداث تكنولوجيا الذكاء الاصطناعي بنجاح في أنشطة الدعم إلى الحد من الوقت المستغرق وزيادة الكفاءة.

وبالإضافة إلى ذلك، فإن تطبيقات الذكاء الاصطناعي في المجال العسكري لديها القدرة على تعزيز تنفيذ القانون الدولي الإنساني والمساعدة في الجهود المبذولة لحماية المدنيين والأعيان المدنية في النزاعات المسلحة. ويمكن أن تكون مفيدة لأنشطة بناء السلام وحفظ السلام، وتعزز قدرات التحقق والرصد لأنظمة تحديد الأسلحة ونزع السلاح وغيرها من أنظمة الامتثال.

كما يطرح الذكاء الاصطناعي في المجال العسكري تحديات غير مسبوقة. فينطوي الذكاء الاصطناعي على مواطن ضعف متأصلة يمكن أن تكون لها عواقب غير مقصودة وتؤدي إلى إضعاف التحكم البشري الفعلي والمسؤولية والمساءلة. ويمكن لاستخدام التعلم العميق أن يجعل من الصعب فهم نماذج الذكاء الاصطناعي وتفسيرها والتنبؤ بها. فعدم الفهم يمكن، على سبيل المثال، أن يجعل ديناميات تصعيد النزاع أكثر غموضاً وأن يزيد من صعوبة التنبؤ بها.

ويجب وضع ضمانات فعالة لضمان احتفاظ البشر بالتحكم والرقابة على نحو فعال فيما يخص تطوير الذكاء الاصطناعي ونشره واستخدامه. ويكتسب هذا الأمر أهمية خاصة كلما كان التطبيق أقرب إلى العمليات القتالية واستخدام القوة، مثل أنظمة دعم اتخاذ القرارات. ويجب أن تظل المساءلة والمسؤولية عن استخدام الذكاء الاصطناعي العسكري وآثاره تقع دائماً على عاتق البشر.

وقد تكون منظومات الذكاء الاصطناعي سريعة التأثر للغاية بجودة بيانات التدريب ومدى تمثيلها. ويمكن أن تؤدي التحيزات المحتملة أو المعلومات المضلِّلة أو المغلوطة أو بيانات التدريب غير المكتملة إلى نماذج تولد نتائج غير دقيقة أو تمييزية. ويمكن أن يتسبب التحيز للتشغيل الآلي في اعتماد المستخدم البشري بشكل مفرط على مخرجات النظام.

وقد تؤدي قدرات الاستجابة الآلية أو الذاتية التشغيل في المجال السيبراني – لا سيما تلك التي لا تحتوي على آليات كافية للتدخل البشري – إلى استجابات غير مقصودة وتصعيد سريع.

كما أن زيادة الاعتماد على التكنولوجيا السيبرانية في المهام التي كانت تُنفذ يدوياً أو بخاصية التشغيل الآلي الأساسية في السابق يصاحبها أيضاً خطر الاستغلال الخبيث لمواطن الضعف الكامنة في تلك التكنولوجيا. ويثير الاعتماد المتزايد على الأنظمة التجارية مخاوف بشأن الاعتماد على مقدمي الخدمات الخارجيين، وفقدان السيطرة على التحديثات وغيرها من مواطن الضعف المتعلقة بالنظم المسجلة الملكية.

وما سبق ذكره مجرد أمثلة على المخاطر المحتملة المرتبطة بتطبيق الذكاء الاصطناعي في المجال العسكري. فهناك أيضًا العديد من المجهولات المجهولة. وفي السياق العسكري، يمكن لهذه العوامل، مجتمعةً أو بمفردها، أن تقوض نتائج المهمة وتشكل مخاطر قانونية وأخلاقية وإنسانية وعسكرية أساسية.

كما تحدد الاستراتيجية النرويجية للذكاء الاصطناعي في قطاع الدفاع (2023) المبادئ الأساسية لتطوير الذكاء الاصطناعي واستخدامه على نحو مسؤول:

- **المشـــروعية** – يجب تطوير تطبيقات الذكاء الاصـــطناعي واســـتخدامها بما يتوافق مع القانون الدولي، بما في ذلك القانون الدولي الإنســـاني والقانون الدولي لحقوق الإنســـان. وعند دراســـة أو تطوير أو حيازة أو اعتماد سلاح أو أداة أو أسلوب للحرب جديد يعتمد على الذكاء الاصطناعي، فإن كل دولة ملزمة بتحديد ما إذا كان اســـتخدامه، أحيانا أو في جميع الأحوال، محظوراً بموجب القانون الدولي لحقوق الإنسان أو أي قاعدة أخرى من قواعد القانون الدولي المنطبقة على الدولة.

- **المســؤولية والمســاءلة** – يجب ضــمان المســؤولية الإنســانية والمســاءلة عن اســتخدام الذكاء الاصطناعي. يجب تحديد سلطة اتخاذ القرار بشأن استخدام نظام الذكاء الاصطناعي والمسؤولية عن استخدامه الفعلي بشكل لا لبس فيه.

- **قابلية التفسير والفهم والتتبع** – يجب أن تكون تطبيقات الذكاء الاصطناعي قابلة للتفسير والفهم والتتبع وأن تتسم بالشفافية بشكل كافٍ.

- **التدريب** – يجب أن يحصـل مشـغلو الذكاء الاصـطناعي على التدريب اللازم لفهم سـلوك تطبيق الذكاء الاصطناعي، بما في ذلك كيفية التعرف على السلوك الشاذ.

- **الموثوقية والســـلامة والأمن** – يجب أن يكون لتطبيقات الذكاء الاصـــطناعي نطاقات اســـتخدام واضـــحة ومحددة جيداً. ويجب أن تخضـــع مرونة تطبيقات الذكاء الاصـــطناعي وموثوقيتها وأمنها للاختبار والتحقق طوال دورتها العمرية الكاملة ضـــمن نطاقات اســـتخدامها. ويجب أن تتمتع تطبيقات الذكاء الاصطناعي بمستويات كافية من الأمن ويجب حمايتها ضد التهديدات الرقمية.

- **التحكم** – يجب ضـمان التحكم البشـري الفعلي. ويجب أن تشـتمل منظومات الذكاء الاصـطناعي على واجهة بينية للتفاعل بين الإنسـان والآلة تكون ملائمة للاسـتخدام المقصـود منها، وتتيح القدرة على تحديد العواقب غير المقصـودة والتخفيف من حدتها، بالإضـافة إلى وسـائل اتخاذ الإجراءات التصحيحية اللازمة إذا كان النظام يعمل بطريقة غير مقصودة.

وثمة حاجة إلى أن يعمق المجتمع الدولي الحوار حول التطبيقات العســكرية للذكاء الاصـــطناعي وآثارها على السـلام والأمن، بما في ذلك التدابير اللازمة لضـمان تحلي الذكاء الاصـطناعي بالمسـؤولية في المجال العســـكري. ويجب إيلاء اهتمام خاص للمنظومات الداعمة للعمليات القتالية، بما في ذلك اســـتخدام الذكاء الاصـــطناعي لأغراض الإلمام بالحالـة ودعم اتخـاذ القرارات، حيث يمكن أن يكون للمخرجـات والسـلوكيات غير المرغوب فيها في تطبيق الذكاء الاصـطناعي وفقدان التحكم البشـري الفعلي عواقب ضـارة بشـكل خاص. وهناك حاجة أيضًـا إلى معالجة الذكاء الاصـطناعي في الحرب الهجينة، بما في ذلك على سـبيل المثال لا الحصـر الذكاء الاصـطناعي في العمليات السـيبرانية، والذكاء الاصـطناعي في الحرب الإلكترونية، والذكاء الاصطناعي في العمليات المعلوماتية.

وتلتزم النرويج بتعزيز التعاون الدولي في مجال تبادل المعلومات وبناء القدرات. ومن خلال تطوير قاعدة معرفية مشـتركة، سـتعمل الدول على تعزيز التوصـل إلى فهم مشـترك وسـد الثغرات وتعزيز الشـفافية وبناء الثقة. وتحقيقاً لهذه الغاية، ستشجِّع النرويج على تطوير ونشر الاستراتيجيات الوطنية ووثائق السياسات

المتعلقة بالتطبيقات العسكرية للذكاء الاصطناعي. وينبغي إيلاء الاهتمام لتدابير الحد من المخاطر وبناء الثقة.

ويمكن أن يساعد تطوير حوكمة دولية ملائمة للذكاء الاصطناعي في الوقت المناسب، مع المرونة اللازمة للاستجابة للتطورات التكنولوجية السريعة، في منع سباقات التسلح القائمة على التكنولوجيا والقيام في الوقت نفسه بضمان أن يدعم الابتكار الأمن العالمي.

باكستان

[الأصل: بالإنكليزية]
[9 نيسان/أبريل 2025]

إن التقدم السريع الذي شهدته تكنولوجيات الذكاء الاصطناعي ودمجها في المجال العسكري من شأنه أن يُحدثا تحولاً جذرياً في الحروب. فيتم دمج الذكاء الاصطناعي بشكل متزايد في العمليات العسكرية من خلال تطبيقاته في منظومات الأسلحة الذاتية التشغيل؛ والقيادة والتحكم؛ وأنظمة دعم اتخاذ القرارات؛ والاستخبارات والمراقبة والاستطلاع؛ والتدريب؛ واللوجستيات؛ والحرب السيبرانية/حرب المعلومات. وفي حين أن هذه التطورات تتيح تحقيق أوجه كفاءة تشغيلية، إلا أنها تشكل أيضًا مخاطر كبيرة على السلام والأمن الدوليين.

التحديات المرتبطة بالذكاء الاصطناعي في المجال العسكري

المخاطر الاستراتيجية: التفاعل مع الأسلحة النووية

ينطوي دمج الذكاء الاصطناعي في منظومات الأسلحة النووية على مخاطر استراتيجية، لا سيما فيما يتعلق بالقيادة والتحكم والاتصالات في المجال النووي. وعندما تدمج قدرات الذكاء الاصطناعي في السياسات المتعلقة بوضع القوة النووية وباستخدامها، فإن ذلك قد ينجم عنه سوء تقدير وحوادث وتترتب عليه عواقب كارثية.

ويعتمد مفهوم الردع النووي اعتمادًا كبيرًا على التعقل والتصور البشريين وعلى اتخاذ القرار السياسي. ومن المحتمل أن يسفر دمج الذكاء الاصطناعي عن إزالة هذه العوامل البشرية الضرورية أو تقليصها بشكل كبير، مما يزيد من مخاطر التصعيد الآلي أو العرضي. وإذ تدرك بعض الدول هذه المخاوف العميقة، فإنها قد التزمت علنًا بالمضي في ممارسة تحكّم بشري فعلي على قرارات استخدام الأسلحة النووية – وهو مبدأ تدعمه باكستان وتحث جميع الدول الحائزة للأسلحة النووية على تأييده.

وفي المناطق التي توجد فيها أسلحة نووية، قد يؤدي الاعتماد على نظم موجهة بالذكاء الاصطناعي لدعم اتخاذ القرارات ومنظومات أسلحة تقليدية كاملة التشغيل الذاتي إلى مخاطر التصعيد. والقضاء التام على التحكم البشري أثناء الأزمات قد يعقّد عملية التحكم في حجم النزاعات ومدتها. وأتمتة الاستجابات في أوضاع متقلبة وعالية المخاطر، لا سيما في المناطق ذات الديناميات النووية المتوترة، قد يضاعف من التداخل بين الأسلحة التقليدية والأسلحة النووية ويؤثر سلبًا على الاستقرار الاستراتيجي.

وقد يخلق استخدام الذكاء الاصطناعي لتقييم البيانات وللاضطلاع بالعمل الاستخباراتي والمراقبة والاستطلاع شعوراً زائفاً بالثقة لدى الدول التي تنوي توجيه ضربات مضادة استباقية لزعزعة الاستقرار أو تستهدف القدرة على توجيه ضربة ثانية، مما يشكل مخاطر جسيمة على الاستقرار الإقليمي والعالمي.

مخاطر التشغيل: فقدان الفاعلية البشرية

من شأن خاصية التشغيل الذاتي القائمة على الذكاء الاصطناعي في العمليات العسكرية أن تحد من الرقابة البشرية، وتصعب من ثم إدارة الأزمات. ولأن وتيرة الحرب تتسارع بـــ "سرعة الآلة"، فإن عملية اتخاذ القرار البشري تصبح مضغوطة بشدة، وهو ما يقلل من فرص التخفيف من حدة الأزمات والتدخل الدبلوماسي.

وقد يولي البشر ثقة مفرطة للتوصيات التي يولّدها الذكاء الاصطناعي انطلاقا من نُظم دعم اتخاذ القرار، حتى لو كانت معيبة أو غير مكتملة، مما يؤدي إلى تحيز التشغيل الآلي. وقد تعتمد القرارات العسكرية الحاسمة بشكل مفرط على مخرجات الآلات، وهو ما قد يدفع القادة إلى تجاهل الحدس البشري أو السياق أو إلى عدم الحذر، ويؤدي بالتالي إلى تصعيد النزاعات عن غير قصد.

وقد تؤدي القدرات المدعومة بالذكاء الاصطناعي، مدفوعة بجاذبية الزيادة في الكفاءة التشغيلية والسباق نحو تحقيق ميزة حاسمة، إلى تعزيز الميل إلى استخدام هذه القدرات، وبالتالي خفض عتبة نشوب النزاع المسلح. وفي أوقات الأزمات، قد يؤدي انخفاض عتبة استخدام القوة إلى زعزعة الاستقرار بشكل كبير.

المخاطر التقنية

قد تنطوي التطبيقات العسكرية للذكاء الاصطناعي على نقاط ضعف تقنية، بما في ذلك التحيز الخوارزمي، وتسميم البيانات، وقابلية التعرض للهجمات السيبرانية. ويمكن أن تنشب النزاعات بسبب خلل في نُظم الإنذار المبكر أو بسبب التلاعب بها أو بسبب هجمات تتمثل في تسميم البيانات. وغالبًا ما تعمل قدرات الذكاء الاصطناعي كـــ "صناديق سوداء"، حيث تنتج قرارات تفتقر إلى الشفافية أو القابلية للتفسير، مما يعقد عملية التحقق والمساءلة. ويمكن أن تؤدّي نقاط الضعف هذه إلى نتائج لا يمكن التنبؤ بها، وإلى أعطال في النظام ومخاطر كبيرة على السلامة التشغيلية. وقدرات الذكاء الاصطناعي التي تُختبر في بيئة تتألف من مجموعات بيانات محددة قد لا تعمل بشكل موثوق في بيئات مختلفة تمامًا مع ديناميات أكثر تعقيداً.

المخاطر المعيارية والقانونية والأخلاقية

يطرح استخدام الذكاء الاصطناعي في المجال العسكري تحديات أخلاقية ومعيارية وقانونية، لا سيما فيما يتعلق بامتثال القانون الدولي الإنساني. فجوهر القانون الدولي الإنساني يعتمد بشكل أساسي على التقدير البشري والسلطة التقديرية وعلى اتخاذ القرارات الحساسة وفقا للسياق – وهي صفات يصعب على منظومات الذكاء الاصطناعي استنساخها. وقد يؤدي تفويض الوظائف الحاسمة مثل اختيار الأهداف والاشتباك، بما في ذلك قرارات استخدام القوة الفتاكة، إلى منظومات ذاتية التشغيل إلى انتهاك مبادئ القانون الدولي الإنساني الأساسية المتمثلة في التمييز والتناسب والتحوط عند الهجوم والضرورة العسكرية.

ومنظومات الذكاء الاصطناعي التي تخلف نتائج لا يمكن التنبؤ بها أو الاعتماد عليها أو تفسيرها تزيد من تعقيد مسألة التقيد بالقانون الدولي الإنساني، وقد تسفر عن أضرار غير قانونية أو غير مقصودة.

وبالإضافة إلى ذلك، فإن غياب القرارات البشرية المباشرة أو الاعتماد المفرط على النُظم الموجهة بالذكاء الاصطناعي لدعم اتخاذ القرارات يثيران أسئلة حرجة حول المساءلة والمسؤولية، مما يجعل الإسناد وتحميل المسؤولية في حالات الأفعال غير القانونية أو غير المشروعة أمراً بالغ الصعوبة. وعند حدوث خطأ ما، فقد يلقي القادة بالمسؤولية على الذكاء الاصطناعي، ويعقدون بالتالي تطبيق مبدأ المساءلة القانونية وفتح التحقيقات المحتملة في جرائم الحرب.

وتنشأ أيضا مخاوف أخلاقية أخرى عن تفويض قرارات الحياة والموت إلى منظومات ذاتية التشغيل، لأن ذلك يقلل من التعاطف والتفكير الأخلاقي والتقدير البشري، وبالتالي يفاقم خطر العنف غير المبرر وسقوط ضحايا مدنيين.

الانتشار والمخاطر الأمنية العالمية

يمثل انتشار تكنولوجيات الذكاء الاصطناعي العسكرية مخاطر كبيرة على الأمن الدولي. وقد يؤدي انتشار قدرات الذكاء الاصطناعي المتقدمة، لا سيما الأسلحة الذاتية التشغيل، إلى خطر إطلاق سباقات تسلح جديدة وزعزعة استقرار البيئات الأمنية الإقليمية والعالمية. كما أن سهولة انتشارها واحتمال أن تستحوذ عليها جهات من غير الدول يزيدان من تفاقم هذه المخاوف.

مقترح الاستجابة الدولية: الدور المركزي لآليات الأمم المتحدة

لتكنولوجيات الذكاء الاصطناعي أغراض عامة، واستخداماتها السلمية جزء لا يتجزأ من تحقيق أهداف التنمية المستدامة. وفي الوقت نفسه، تكون الآثار المترتبة على الذكاء الاصطناعي في المجال العسكري شاملة لعدة قطاعات ويمكن أن تؤثر بشكل كبير على السلم والأمن الدوليين، مما يستلزم استجابة دولية منسقة.

وتعترف باكستان بقيمة مبادرات حوكمة استخدام الذكاء الاصطناعي خارج الأمم المتحدة، لكنّها تظل تدرك أوجه قصورها، لا سيما فيما يتعلق بالمشاركة العالمية والشرعية الرسمية المتعددة الأطراف. وفي حين يمكن لهذه المبادرات أن تكمل جهود الأمم المتحدة من خلال تعزيز الحوار والإرادة السياسية، قد يؤدّي تنفيذها بمعزل عن بعضها البعض إلى حالة تشظ. ولذلك، ينبغي إجراء المناقشات حول التطبيقات العسكرية للذكاء الاصطناعي داخل الأمم المتحدة لضمان الشمولية والشرعية وإطار عالمي متماسك يعكس مصالح جميع الدول.

ولهذه الأسباب، يجب أن تظل الأمم المتحدة محور أي استجابة دولية. ويجب أن تلعب آليات الأمم المتحدة لنزع السلاح دورا جوهريا في وضع إطار حوكمة دولي للذكاء الاصطناعي العسكري ومنع تشظي المشهد المعياري. ويقتضي حجم الآثار العسكرية للذكاء الاصطناعي وطابعها الجديد توفير استجابة متعددة الأوجه وشاملة ومتعددة الأطراف. والعضوية العالمية للأمم المتحدة يبوئها مكانة فريدة باعتبارها المنتدى المثالي الذي يسمح لجميع الدول – المتقدمة والنامية على حد سواء – بإسماع صوتها.

ولن يكفي منتدى واحد أو أداة واحدة. بل هناك حاجة إلى استراتيجية منظمة تستخدم هيئات متعددة لنزع السلاح تابعة للأمم المتحدة، على أن يعالج كل منتدى القضية من زاويته الفريدة وفي إطار ولايته، معالجة متكاملة. ونقترح الاستفادة من جميع المنتديات ذات الصلة، من الجمعية العامة ولجنتها الأولى إلى هيئة نزع السلاح ومؤتمر نزع السلاح واتفاقية حظر أو تقييد استعمال أسلحة تقليدية معينة يمكن اعتبارها مفرطة الضرر أو عشوائية الأثر. ومن شأن هذا النهج أن يعالج الأبعاد الاستراتيجية والإنسانية والقانونية والتقنية بشكل شامل، مع تجنّب الثغرات والتكرار. وينبغي أن يكون عمل كل منتدى منها مفيداً للبقية، عن طريق خلق تآزر يحقق الهدف المشترك المتمثل في الجمع بين التخفيف من مخاطر الذكاء الاصطناعي العسكري والحفاظ على الاستخدام السلمي للذكاء الاصطناعي.

مؤتمر نزع السلاح

ينبغي أن يولي مؤتمر نزع السلاح الأولوية لمعالجة المخاطر الاستراتيجية المرتبطة بالذكاء الاصطناعي العسكري، لا سيما في المجال النووي، بما يتماشى مباشرة مع البندين 1 و 2 من جدول أعماله (”وقف سباق التسلح النووي ونزع السلاح النووي“ و ”منع نشوب حرب نووية، بما في ذلك جميع المسائل ذات الصلة“). وفي عام 2023، اقترحت باكستان إدراج بند جديد في جدول أعمال مؤتمر نزع السلاح بشأن هذا الموضوع (CD/2334).

وفي إطار هذا البند الجديد من جدول الأعمال، ينبغي لمؤتمر نزع السلاح أن ينشئ هيئة فرعية أو فريقاً مخصصاً مكلفاً على وجه التحديد بدراسة مخاطر الذكاء الاصطناعي العسكري على الاستقرار، وتقييم كيفية مساهمة الذكاء الصناعي في المخاطر النووية، ومتابعة المفاوضات بشأن تدابير ملموسة. ويمكن أن تشمل هذه التدابير ما يلي:

- الالتزام بالحفاظ على التحكم البشري، وعدم الاستغناء عن التقدير البشري فيما يتعلق بقرارات استخدام الأسلحة النووية؛
- حظر استخدام قدرات الذكاء الاصطناعي للتلاعب بالبيانات أو استهداف أنظمة القيادة والتحكم والاتصالات في المجال النووي؛
- وضع تدابير ضبط النفس فيما يتعلق بنشر واستخدام بعض قدرات الذكاء الاصطناعي التي يمكن أن تؤدي إلى توجيه ضربات استباقية وتساهم في تصاعد المخاطر النووية.

ويعد مؤتمر نزع السلاح مناسباً بشكل فريد لهذه المناقشات، لأنه يجمع كل الدول القوية عسكرياً على قدم المساواة ويعمل بتوافق الآراء، وبالتالي يحمي المصالح الأمنية الحيوية لجميع الدول. ومعالجة هذه القضية قد يعيد تنشيط عمل مؤتمر نزع السلاح، بما يدلّ على استجابة المؤتمر للتهديدات الجديدة والناشئة.

هيئة نزع السلاح

توجد هيئة نزع السلاح التابعة للأمم المتحدة، بفضل عضويتها العالمية وولايتها التداولية، في وضع مثالي لوضع مبادئ توجيهية وتوصيات عملية بشأن الاستخدام العسكري المسؤول للذكاء الاصطناعي. وعبر التاريخ، وضعت هيئة نزع السلاح التابعة للأمم المتحدة مبادئ توجيهية مماثلة (على سبيل المثال، تدابير بناء الثقة في عام 1988 والنُّهج الإقليمية لنزع السلاح في عام 1993).

وفي إطار فريقها العامل الثاني، يمكن أن تضع هيئة نزع السلاح مبادئ توجيهية وتوصيات بشأن تدابير بناء الثقة والأمن المتعلقة بتطبيقات الذكاء الاصطناعي العسكرية على المستويين العالمي والإقليمي. وقد تشمل العناصر الرئيسية إعادة التأكيد على الأسس المعيارية، والتوصية بتدابير للتخفيف من المخاطر التشغيلية والتقنية، ووضع استراتيجيات عسكرية للحد من مخاطر الذكاء الاصطناعي، ومراعاة مخاوف الانتشار مع ضمان الوصول العادل إلى الاستخدامات السلمية للذكاء الاصطناعي.

اللجنة الأولى للجمعية العامة للأمم المتحدة

ينبغي للجنة الأولى للجمعية العامة للأمم المتحدة أن تضفي طابعا مؤسسيا على عملية تقديم الأمين العام للأمم المتحدة تقارير تقييم منتظمة، وأن تحتفظ بفهرس للتطور التكنولوجي لقدرات الذكاء الاصطناعي العسكرية والمخاطر المرتبطة بها استنادًا إلى المعلومات التي تتبادلها الدول الأعضاء بشكل طوعي. ومن شأن هذه التقييمات الدورية أن توفر رؤى موثوقة حول تطور القدرات، مما يوفر معلومات في الوقت المناسب ويسهل وضع استجابات سياساتية دولية مستنيرة.

ويمكن للجنة الأولى، عند استعراضها لهذه التقارير، أن تعقِد مناقشات مخصصة حول الذكاء الاصطناعي وأن تقوم ربما بإنشاء فريق عامل مفتوح العضوية في إطار الجمعية العامة، للتفاوض، إذا لزم الأمر، بشأن منصّة أكثر مؤسسية مثل سجلّ الأمم المتحدة للتطبيقات العسكرية للذكاء الاصطناعي (على الرغم من أنه يفضَّل في الوقت الحالي الاستفادة من المنتديات الموجودة).

ويمكن أن تحدد هذه التقارير أيضا المجالات التي يظهر فيها توافق في الآراء أو التي تحتاج إلى مزيد من العمل، لتوجيه جداول أعمال منتديات مثل مؤتمر نزع السلاح وهيئة نزع السلاح والاتفاقية المتعلقة بأسلحة تقليدية معينة.

الاتفاقية المتعلقة بأسلحة تقليدية معينة

لا يزال فريق الخبراء الحكوميين المنشأ بموجب الاتفاقية المتعلقة بأسلحة تقليدية معينة ضرورياً لمعالجة الآثار الإنسانية والأخلاقية والقانونية المترتبة على منظومات الأسلحة الفتاكة الذاتية التشغيل. كما أن طبيعته الشاملة (إشراك المجتمع المدني واللجنة الدولية للصليب الأحمر كمراقبين) تعد ميزة.

وتقدر باكستان العمل الذي أنجزه فريق الخبراء الحكوميين منذ عام 2017، ولا سيما المبادئ التوجيهية الأحد عشر التي وُضعت في عام 2019. ومع ذلك، كان التقدم المحرز في إطار الاتفاقية بطيئاً وقائماً إلى حدّ كبير على مبادئ وليس على لوائح تنظيمية ملموسة. وتتفق باكستان مع التقييمات التي تفيد بأن المناقشات التي أجريت في إطار الاتفاقية أولت "اهتماماً غير كاف ومتراجعاً" للأبعاد الأمنية للأسلحة التي تعمل بالذكاء الاصطناعي، مما يبرز الحاجة إلى اتخاذ إجراءات تكميلية داخل مؤتمر نزع السلاح وغيره من المنتديات. ومع ذلك، وعلى صعيد العمل الإنساني، ينبغي على فريق الخبراء الحكوميين المنشأ بموجب الاتفاقية مواصلة وتكثيف عمله.

وتدعو باكستان إلى اختتام المفاوضات بشأن بروتوكول للاتفاقية المتعلقة بأسلحة تقليدية معينة يكون ملزم قانوناً ويحظر منظومة الأسلحة الفتاكة الذاتية التشغيل التي تعمل دون رقابة بشرية أو لا يمكنها امتثال القانون الدولي الإنساني. وتسمح الولاية الحالية لفريق الخبراء الحكوميين للدول الأعضاء بأن تضع

عناصر مثل هذا الصك بغرض عرضها في المؤتمر الاستعراضي السابع للأطراف المتعاقدة السامية في الاتفاقية، وبأن تباشر ربما بعد ذلك مفاوضات رسمية.

الخلاصة

تشدد باكستان على ضرورة اتخاذ إجراءات دولية منسقة وشاملة للتخفيف من المخاطر الكبيرة للذكاء الاصطناعي العسكري. وهي تتصور نهجاً للحوكمة يحقق توازناً بين الأمن والتنمية، ويضمن الاستقرار مع السماح بتطوير الذكاء الاصطناعي على نحو مفيد. وبفضل استراتيجية منظمة ومتعددة المنتديات داخل الأمم المتحدة، يمكن للمجتمع الدولي أن يضع حواجز معيارية قوية تحمي الأمن الدولي وتحافظ على الوصول العادل وغير التمييزي إلى الاستخدامات السلمية للذكاء الاصطناعي.

بيرو

[الأصل: بالإسبانية]
[11 نيسان/أبريل 2025]

تطلب الجمعية العامة إلى الأمين العام، في الفقرة 7 من قرارها 239/79 الذي اتخذته في 24 كانون الأول/ديسمبر 2024 وصوتت بيرو لصالحه، أن يلتمس:

آراء الدول الأعضاء والدول المراقبة بشأن الفرص والتحديات التي يطرحها تطبيق الذكاء الاصطناعي في المجال العسكري على السلام والأمن الدوليين، مع التركيز بوجه خاص على مجالات غير منظومات الأسلحة الفتاكة الذاتية التشغيل، وأن يقدم تقريرا موضوعيا يلخص تلك الآراء ويصنف المقترحات المعيارية القائمة والناشئة، ويحتوي على مرفق يتضمن هذه الآراء، إلى الجمعية العامة في دورتها الثمانين لكي تواصل الدول مناقشتها.

وفي هذا الصدد، تعرض بيرو أدناه بعض جوانب موقفها بهدف المساهمة في إعداد تقرير الأمين العام المذكور أعلاه.

أولا - أهمية الذكاء الاصطناعي في المجال العسكري

تقر بيرو بالتطور السريع والديناميكي للتكنولوجيات الناشئة في المجال العسكري، ولا سيما التطبيقات المحتملة للذكاء الاصطناعي. وتراقب بيرو عن كثب التطورات في هذا المجال - بما في ذلك الطرق التي يبدو أن الذكاء الاصطناعي يحدث من خلالها تحولا في العمليات العسكرية، من استخدام الطائرات المسيّرة الذاتية التشغيل إلى نظم دعم اتخاذ القرارات - وترى أن من الأهمية بمكان التشجيع على وجود حوار مستمر ومتعدد الأطراف يهدف إلى وضع مبادئ تضمن الاستخدام الأخلاقي والمسؤول لهذه الأدوات.

ونظرًا لإمكانية إدماج الذكاء الاصطناعي في منظومات الأسلحة ونظم دعم العمليات العسكرية على حد سواء، ترى بيرو أن من الضروري معالجة التحديات والشواغل التي يثيرها استخدام الذكاء الاصطناعي من المنظور الإنساني والقانوني والأمني والتكنولوجي والأخلاقي، بما في ذلك المخاطر المرتبطة بأشكال التحيز الخوارزمية. وتتفاقم هذه الشواغل بسبب الآثار المحتملة لاستخدام هذه التكنولوجيا على الاستقرار والأمن الدوليين.

ويثير هذا الأمر قلقًا أكبر نظرًا للآثار المترتبة على استخدام الذكاء الاصطناعي على صعيد الأسلحة النووية وغيرها من أسلحة الدمار الشامل. ولذلك يجب التأكيد على مبدأ التحكم البشري الفعلي.

ثانيا - الآراء

الامتثال للقانون الدولي

يجب أن يكون تطوير التكنولوجيات القائمة على الذكاء الاصطناعي وتطبيقها واستخدامها في المجال العسكري متوافقا مع القانون الدولي، بما في ذلك القانون الدولي لحقوق الإنسان والقانون الدولي الإنساني، وكذلك مع المبادئ الأساسية المنصوص عليها في ميثاق الأمم المتحدة.

وفي هذا الصدد، فإن أي أنشطة معيارية لتنظيم الذكاء الاصطناعي في المجال العسكري يجب أن تضمن استخدامه المسؤول والأخلاقي، وأن تضمن أيضًا عدم انتشار التكنولوجيات العسكرية القائمة على الذكاء الاصطناعي والوصول العادل إلى المعارف والقدرات التكنولوجية.

والغرض من ذلك ضمان أن يكون أي استخدام للذكاء الاصطناعي يحترم الكرامة الإنسانية ويحمي السكان المدنيين ويضمن الاستقرار والسلام الدوليين.

التعرف على الفوائد والمخاطر

يوفر الذكاء الاصطناعي فرصاً قيّمة لفهم أفضل للأوضاع العملياتية وبالتالي تحسين تنفيذ القانون الدولي الإنساني وحماية المدنيين والأعيان المدنية.

ومع ذلك، قد ينطوي استخدامه على مخاطر متوقعة وأخرى غير متوقعة في المجال العسكري، ومنها على سبيل المثال تلك الناشئة عن أشكال التحيز الخوارزمية أو عيوب التصميم أو سوء الاستخدام أو الاستخدام الخبيث. وعلاوة على ذلك، يمكن أن يكون للذكاء الاصطناعي أثر على الديناميات الإقليمية والعالمية المعقدة نظراً لأنه يؤثر على مخاطر التصعيد وسوء التقدير وخفض عتبة النزاع ونشوء سباق التسلح.

التنمية المسؤولة

يجب أن يعزز استخدام الذكاء الاصطناعي في المجال العسكري السلام وحماية السكان المدنيين ومفهوم أن التقدم التكنولوجي يجب أن يكون مكملاً للقدرات البشرية وليس بديلاً عنها.

وتماشياً مع المبادئ المطبقة على منظومات الأسلحة الذاتية التشغيل، يجب أن يضمن تطبيق الذكاء الاصطناعي في المجال العسكري ألا تنقل المسؤولية والمساءلة إلى الآلات أبدا. وفي هذا الصدد، تشدد بيرو على ضرورة الحفاظ على التحكم البشري الفعلي في جميع القرارات التي تنطوي على استخدام القوة.

ويجب معالجة جميع المخاطر والتحديات المتعلقة بهذه التكنولوجيا بطريقة شاملة طوال دورة حياتها.

ويمكن وضع ضوابط وضمانات لمنع إساءة استخدام هذه التكنولوجيا في المجال العسكري دون إعاقة أنشطة البحث والتطوير والتجارب والابتكار المتعلقة بالذكاء الاصطناعي في المجالات الأخرى.

التنفيذ والشفافية

إن تحديد الاستراتيجيات والمبادئ والمعايير والقواعد، فضلاً عن السياسات الوطنية والأطر القانونية، التي تضمن الاستخدام المسؤول للذكاء الاصطناعي في المجال العسكري مسألة ذات أولوية.

ومن المهم أيضاً وضع تدابير لبناء الثقة والحد من المخاطر، فضلاً عن آليات لتبادل الممارسات الجيدة، من أجل تحقيق الشفافية والتعاون بين الدول.

شكل المناقشات

ترى بيرو أن من الضروري الحفاظ على حوار مستمر على الصعيدين العالمي والإقليمي وفيما بين الدول بشأن وضع تدابير لضمان توظيف الذكاء الاصطناعي على نحو مسؤول في المجال العسكري.

وتدعو أيضا إلى المشاركة الشاملة في هذا المجال، حيث يتم النظر في آراء الدول، ولا سيما الدول النامية، وكذلك مساهمات أصحاب المصلحة الآخرين، مثل الصناعة والأوساط الأكاديمية والمجتمع المدني والمنظمات الإقليمية والدولية.

ومن المهم مراعاة أن الدول والمناطق المختلفة تقف في مراحل مختلفة من إدماج قدرات الذكاء الاصطناعي في المجال العسكري وتعمل في بيئات أمنية مختلفة.

وهذا يؤكد على قيمة تعزيز بناء القدرات في البلدان النامية وتعزيز التعاون الدولي بهدف تقليص الفجوات القائمة وتعزيز مشاركة هذه البلدان في المناقشات حول استخدام هذه التكنولوجيا.

مشاركة بيرو في المناقشات الدولية

مؤتمرات قمة تسخير الذكاء الاصطناعي المسؤول في المجال العسكري

- شاركت بيرو في مؤتمري القمة لعامي 2023 و 2024 وفي حلقة العمل الإقليمية ذات الصلة
- المصادقة على الإعلان الختامي لمؤتمر قمة تسخير الذكاء الاصطناعي المسؤول في المجال العسكري لعام 2024 ("مخطط العمل")

الإعلان السياسي بشأن الاستخدام العسكري المسؤول للذكاء الاصطناعي وخاصية التشغيل الذاتي:

- شاركت بيرو في الجلسة العامة الافتتاحية بشأن هذه المبادرة بصفة مراقب، ثم أضفت الطابع الرسمي على تأييدها

مؤتمر القمة المعني بالإجراءات المتخذة على صعيد الذكاء الاصطناعي – المحادثات العسكرية (باريس، 2025)

- شاركت بيرو مشاركة رفيعة المستوى ووقعت على إعلان باريس بشأن الحفاظ على التحكم البشري في منظومات الأسلحة التي تعتمد على الذكاء الاصطناعي.

جمهورية كوريا

[الأصل: بالإنكليزية]
[11 نيسان/أبريل 2025]

يحمل الذكاء الاصـــطناعي في طياته، باعتباره تكنولوجيا تمكينية، إمكانات تؤهلها لإحداث تحول جذري في أبعاد متعددة من الشؤون العسكرية – بدءًا من صنع القرار وجمع المعلومات الاستخباراتية ووصولا إلى الخدمات اللوجستية والمراقبة ومنظومات القيادة والتحكم. ومع تسارع وتيرة تطور الذكاء الاصطناعي، يتزايد اهتمام الدول بتسخير هذه التكنولوجيا في المجال العسكري.

وإذ يتزايد تغلغل قدرات الذكاء الاصطناعي والنظم القائمة على الذكاء الاصطناعي في العمليات العسكرية، فإنها تنطوي على فرص وتحديات على حد سواء، لا سيما فيما يخص السلام والأمن الدوليين. وتثير هذه التطورات أسئلة مهمة من المنظور الإنساني والقانوني والأمني والتكنولوجي والأخلاقي.

ولأغراض هذه المساهمة، تركز الآراء الواردة أدناه بشكل خاص على مجالات أخرى غير منظومات الأسلحة الفتاكة الذاتية التشغيل.

الفرص التي ينطوي عليها الذكاء الاصطناعي في المجال العسكري

إن قدرات الذكاء الاصـــطناعي والمنظومات المدمجة بالذكاء الاصـــطناعي، بما في ذلك تلك المستخدمة في نظم الاستخبارات والمراقبة والاستطلاع ونظم دعم القرار، تمكّن من زيادة الإلمام بالحالة، وتعزيز الدقة، وتحسين الكفاءة عبر معالجة بيانات واسعة النطاق ودعم التحسين إلى المستوى الأمثل، وتوليد رؤى تنبؤية. ويمكن لهذه القدرات والمنظومات أن تساهم في الحفاظ على السلام والأمن الدوليين وتعزيزهما.

1 - *تعزيزُ تنفيذ القانون الدولي الإنساني والمساعدةُ في حماية المدنيين والأعيان المدنية في النزاعات المسلحة*

يمكن لنظم الاســـتخبارات والمراقبة والاســـتطلاع ونظم دعم اتخاذ القرار القائمة على الذكاء الاصطناعي أن تعزز تنفيذ المبادئ الأساسية للقانون الدولي الإنساني – أي التمييز والتناسب والحيطة في الهجوم – من خلال التمكين من إجراء تقييمات أكثر دقة في ساحة المعركة وتحسين الإلمام بالحالة. ويمكن للذكاء الاصـــطناعي أن يســـاعد على التمييز بين المحاربين وغير المحاربين، وتقييم الأضـــرار الجانبية المحتملة، باستخدام معلومات مستنيرة حسنة التوقيت. ويساعد الذكاء الاصطناعي في تحديد متى يكون اتخاذ تدابير احترازية لحماية المدنيين والبنية التحتية المدنية ضروريا وما مستواه الملائم، من خلال تحسين الإلمام بما يدور في ساحة المعركة، بما في ذلك وجود المدنيين.

2 - *دعم عمليات حفظ السلام*

يمكن للذكاء الاصطناعي أن يدعم رصد اتفاقات وقف إطلاق النار واتفاقات السلام. ويمكن له أيضـــا أن ييسر آليات الإنذار المبكر من أجل الكشف عن الانتهاكات المحتملة، وهو ما سيعزز فعالية البعثات وسلامتها. وقد أطلقت جمهورية كوريا مشروعًا تجريبيًا للمعسكرات الذكية في وحدة هانبيت التابعة لبعثة الأمم المتحدة في جنوب السودان بهدف تعزيز السلامة والكفاءة والقدرات التشغيلية لمعسكرات الأمم المتحدة لحفظ السلام من خلال تطبيق الذكاء الاصطناعي وغيره من التكنولوجيات الناشئة.

3 - تعزيز قدرات التحقق والرصد فيما يتعلق بنظم تحديد الأسلحة والامتثال

يمكن للذكاء الاصطناعي أن يعزز قدرات آليات التحقق الدولية على رصد الامتثال لاتفاقات تحديد الأسلحة وعدم الانتشار. ويمكن للوكالة الدولية للطاقة الذرية أن تستفيد من الذكاء الاصطناعي لزيادة كفاءة العمليات المتصلة بالضمانات، لا سيما تلك التي تنطوي على تصنيف البيانات والعثور على أنماط والتعرف على القيم المتطرفة في البيانات. ويمكن للمنظومات القائمة على الذكاء الاصطناعي أن تساعد أيضًا على تحديد المؤشرات المبكرة لاستخدام أسلحة كيميائية أو بيولوجية والكشف عن أساليب التهرب من الجزاءات التي لا تنفك تزداد تطورًا، وهو ما سيعزز النظم الدولية لعدم الانتشار.

وبالإضافة إلى الفرص الموضحة أعلاه، يمكن للذكاء الاصطناعي أن يساعد على التخفيف من حدة المخاطر الاستراتيجية – كسوء التقدير وسوء التفاهم والتصعيد غير المقصود – من خلال تحسين تحليل سلوك الجهات الفاعلة وتعزيز القدرة على كشفه والاستجابة له بشكل استباقي. وعلاوة على ذلك، يمكن لقدرات الذكاء الاصطناعي أن تيسر تطوير القدرات اللازمة لتعزيز الموقف الدفاعي في المجال السيبراني، وحماية البنية التحتية الوطنية الحيوية، ومكافحة الإرهاب، ضمن أمور أخرى.

التحديات التي يطرحها الذكاء الاصطناعي في المجال العسكري

إن تطبيق الذكاء الاصطناعي في الميدان العسكري قد يؤدي إلى ظهور تحديات جديدة أو إلى تفاقم التحديات القائمة ما لم يتم تطويره ونشره واستخدامه بشكل مسؤول.

وقد تنشأ التحديات من الخصائص التقنية والتشغيلية للذكاء الاصطناعي. فعلى سبيل المثال، طابعه المبهم الذي يشبه الصندوق الأسود يجعل من الصعب فهم كيفية وسبب توليده لمخرجات معينة، مما يؤدي إلى الحد من قابليتها للتفسير والتتبع. ويمكن لعيوب التصميم والتحيزات غير المقصودة في البيانات أو الخوارزميات أو بنية المنظومة أن تفضي إلى أعطال أو مخرجات تنحرف عن الأهداف المقصودة. وقد يؤدي الاعتماد المفرط على منظومة الذكاء الاصطناعي، وكذلك تحيز الآلة أو عدم كفاية التدريب، إلى إثارة مشاكل تعزى إلى عدم وجود مستوى ملائم من التقدير البشري والمشاركة البشرية. ويمكن لهذه العوامل أن تزيد من احتمال حدوث سوء التقدير أو سوء التفسير أو التصعيد غير المقصود في النزاعات، مما يشكل تحدياً أمام السلام والأمن الدوليين.

ويمكن لطابع الاستخدام المزدوج الذي تتسم به تكنولوجيات الذكاء الاصطناعي أن يؤدي إلى زيادة مخاطر إساءة استخدامها أو التعسف في استخدامها من قبل جهات فاعلة غير مسؤولة ذات نوايا خبيثة. فعلى سبيل المثال، في المجال السيبراني، قد تتسارع حملات المعلومات المضللة والهجمات السيبرانية القائمة على الذكاء الاصطناعي، مثل تسميم البيانات والانتحال. وعلاوة على ذلك، قد تستغل جهات فاعلة غير مسؤولة تكنولوجيات الذكاء الاصطناعي لتيسير تطوير أسلحة كيميائية أو بيولوجية جديدة، وهو ما يثير مخاوف الانتشار ويزيد من المخاطر التي تهدد السلام والأمن الدوليين.

تطبيق الذكاء الاصطناعي المسؤول في المجال العسكري

لكي يتم جني فوائد الذكاء الاصطناعي والاستفادة من الفرص التي يتيحها مع التصدي في الوقت نفسه للمخاطر والتحديات المرتبطة به، لا بد من التحلي بالمسؤولية في تطوير ونشر واستخدام قدرات الذكاء الاصطناعي والمنظومات القائمة عليه في المجال العسكري على مدار دورة حياتها بأكملها.

وتلتزم جمهورية كوريا بضمان وتعزيز التطبيق المسؤول للذكاء الاصطناعي في المجال العسكري. ويشمل ذلك المبادئ والتدابير الرئيسية التالية:

- ينبغي أن يكون الذكاء الاصطناعي أخلاقيًا ومتمحورًا حول الإنسان.
- يجب أن تُطبق قدرات الذكاء الاصطناعي في المجال العسكري وفقاً للقانون الدولي الساري، بما في ذلك القانون الدولي الإنساني والقانون الدولي لحقوق الإنسان.
- يظل البشر مسؤولين وخاضعين للمساءلة عن استخدام الذكاء الاصطناعي في المجال العسكري وآثار تطبيقاته فيه، ولا يمكن أبدا نقل هذه المسؤولية والمساءلة إلى الآلات.
- يجب ضمان موثوقية تطبيقات الذكاء الاصطناعي وجدارتها بالثقة من خلال وضع الضمانات المناسبة للحد من مخاطر الأعطال أو العواقب غير المقصودة، بما في ذلك تلك الناجمة عن البيانات والخوارزميات وغيرها من التحيزات.
- يجب الحفاظ على مستوى ملائم من المشاركة البشرية في تطوير الذكاء الاصطناعي ونشره واستخدامه في المجال العسكري، بما في ذلك اتخاذ التدابير المناسبة المتصلة بالتقدير والتحكّم البشريين في استخدام القوة.
- ينبغي أن يكون الموظفون المعنيون قادرين على فهمِ المخرجات التي تنتجها قدرات الذكاء الاصطناعي في المجال العسكري، بما في ذلك المنظومات القائمة على الذكاء الاصطناعي، وعلى تفسيرها وتتبعها والثقة بها. ويجب أن تستمر الجهود المبذولة لتحسين قابلية الذكاء الاصطناعي في المجال العسكري للتفسير والتتبع.

وتؤيد جمهورية كوريا إجراء مناقشات وحوارات بشأن اتخاذ مزيد من التدابير لضمان تحلي الذكاء الاصطناعي في المجال العسكري بالمسؤولية، بسبل منها وضع أطر معيارية دولية؛ وبروتوكولات اختبار وتقييم صارمة؛ وعمليات شاملة للتحقق والمصادقة والاعتماد؛ وآليات وطنية قوية للرقابة؛ وعمليات رصد مستمرة؛ وبرامج وتمارين تدريبية شاملة؛ وتعزيز الأمن السيبراني؛ وأطر واضحة للمساءلة.

ويشكل وضع تدابير رقابية وأمنية قوية أمرًا بالغ الأهمية لمنع الجهات الفاعلة غير المسؤولة من اقتناء وإساءة استخدام قدرات الذكاء الاصطناعي التي قد تكون ضارة في المجال العسكري، بما في ذلك المنظومات القائمة على الذكاء الاصطناعي.

وتشجع جمهورية كوريا على وضع تدابير فعالة لبناء الثقة واتخاذ التدابير المناسبة للحد من المخاطر، فضلا عن تبادل المعلومات بين الدول وإجراء مشاورات فيما بينها بشأن الممارسات الجيدة والدروس المستفادة.

وتشدد جمهورية كوريا على ضرورة منع الجهات الفاعلة من الدول وغير الدول من استخدام قدرات الذكاء الاصطناعي في المساهمة في انتشار أسلحة الدمار الشامل، وتؤكد على أن قدرات الذكاء الاصطناعي ينبغي ألا تعرقل جهود تحديد الأسلحة ونزع السلاح وعدم الانتشار. ومن الأهمية بمكان الحفاظ على التحكم البشري والدور البشري في جميع الإجراءات الحاسمة في توجيه وتنفيذ القرارات السيادية المتعلّقة باستخدام الأسلحة النووية، دون المساس بالهدف النهائي المتمثل في إيجاد عالم خالٍ من الأسلحة النووية.

وينبغي تطوير قدرات الذكاء الاصطناعي والمنظومات القائمة على الذكاء الاصطناعي في المجال العسكري ونشرها واستخدامها بطريقة تحافظ على السلام والأمن الدوليين ولا تعيقهما.

الحوكمة المقبلة لاستخدام الذكاء الاصطناعي في المجال العسكري

عند وضع تصور للحوكمة المقبلة لاستخدام الذكاء الاصطناعي في المجال العسكري، من الأهمية بمكان تعزيز فهم موحَّد لتكنولوجيا الذكاء الاصطناعي – قدراتها وحدودها – وفهم مشترك لتطبيقات الذكاء الاصطناعي الممكنة في المجال العسكري، وكذلك آثاره على السلام والأمن الدوليين.

ومن المهم أيضًا بناء القدرات، لا سيما بالنسبة إلى البلدان النامية، لتعزيز مشاركتها الكاملة في مناقشات الحوكمة وتيسير اتباع نهج مسؤول وإيجاد فهم مشترك فيما يتعلق بتطوير ونشر واستخدام الذكاء الاصطناعي في المجال العسكري. ويمكن لتبادل المعارف والممارسات الجيدة والدروس المستفادة أن ييسر أيضا التوصل إلى فهم مشترك.

ونظراً للوتيرة السريعة التي يتقدم بها الذكاء الاصطناعي، ينبغي أن تكون آليات الحوكمة مرنة بما يكفي للتكيف مع هذا التقدم. وتدعم جمهورية كوريا أيضا اتباع نهج متوازن يغتنم الفرص ويتصدى للمخاطر على حد سواء. وقد تؤدي المحادثات المتعلقة بالحوكمة التي تركز بشكل مفرط على المخاطر أو التي تكون تقييدية أكثر من اللازم إلى خنق الابتكار وحجب إمكانات الذكاء الاصطناعي في دعم السلام والأمن الدوليين. وينبغي ألا تكون الحوكمة المقبلة عائقًا أمام الابتكار، بل ينبغي أن تدعمه وتؤدي دور عامل تمكيني للتطبيق المسؤول للذكاء الاصطناعي في المجال العسكري.

وبما أن المجتمع الدولي لا يزال في المراحل الأولى من فهم آثار الذكاء الاصطناعي في المجال العسكري على السلام والأمن الدوليين، وبالنظر إلى الحالة الراهنة الذي وصل إليه التطور التكنولوجي والسياساتي، سيكون من السابق لأوانه تحديد مسار ضيق لحوكمة الذكاء الاصطناعي أو وضع صكوك أو معايير ملزمة قانونًا دون التوصل إلى فهم موحَّد ومشترك لما هو الذكاء الاصطناعي المسؤول في المجال العسكري. وتعتقد جمهورية كوريا أن المناقشات المتعلقة بالحوكمة ينبغي أن تكون واقعية وأن تسير بشكل تدريجي، مسترشدة بالحوار المستمر.

وجمهورية كوريا، إذ تدرك أن الابتكار في مجال الذكاء الاصطناعي يقوده حاليا القطاع الخاص، تعتقد أن جهود الحوكمة المستقبلية يجب أن تتبنى نهجًا مفتوحًا وشاملًا للجميع يشارك فيه أصحاب المصلحة المتعددون، بما في ذلك القطاع الصناعي والأوساط الأكاديمية والمجتمع المدني والمنظمات الإقليمية والدولية.

وتنوه جمهورية كوريا بالجهود الوطنية والإقليمية والعالمية المبذولة لاغتنام فرص الذكاء الاصطناعي في المجال العسكري والتصدي لتحدياته، بما في ذلك وضع الاستراتيجيات والتشريعات والمبادئ والمعايير والسياسات والتدابير الوطنية ذات الصلة، وتسلّم بأهمية تعزيز الحوار على جميع المستويات.

ولضمان التطبيق المسؤول للذكاء الاصطناعي في المجال العسكري، أنشأت جمهورية كوريا حديثًا شعبة سياسة البيانات والفريق المعني بسياسة الذكاء الاصطناعي في مجال الدفاع داخل وزارة الدفاع الوطني في عامي 2022 و 2025 على التوالي. وفي عام 2024، أطلقت الوزارة لجنة البيانات المتعلقة بالدفاع والذكاء الاصطناعي لكي تكون أعلى هيئة للتداول وصنع القرار.

وبغية تعزيز الحوار، استضافت جمهورية كوريا، بالتعاون مع سنغافورة، وكينيا، والمملكة المتحدة، وهولندا، القمة الثانية لتسخير الذكاء الاصطناعي المسؤول في المجال العسكري في أيلول/سبتمبر 2024 في سيول. وقد كانت قمم تسخير الذكاء الاصطناعي المسؤول في المجال العسكري وسلسلة من المشاورات الإقليمية بشأن الذكاء الاصطناعي المسؤول في المجال العسكري في عام 2024 بمثابة حاضنة لتبادل الخبرات وتعزيز الحوار الشامل وتعزيز التفاهم المتبادل. واستشرافًا للمستقبل، ستواصل القمة الثالثة لتسخير الذكاء الاصطناعي المسؤول في المجال العسكري، التي ستُعقد في إسبانيا في أيلول/سبتمبر 2025، إلى جانب المشاورات الإقليمية القادمة بشأن الذكاء الاصطناعي المسؤول في المجال العسكري في عام 2025، توجيهَ جهود المجتمع الدولي نحو التطبيق المسؤول للذكاء الاصطناعي في المجال العسكري.

وتعتقد جمهورية كوريا أن المناقشات في إطار الأمم المتحدة، بما في ذلك اللجنة الأولى للجمعية العامة وهيئة نزع السلاح، بشأن التطبيق المسؤول للذكاء الاصطناعي في المجال العسكري، ينبغي أن تعمل بطريقة تكاملية مع المبادرات الأخرى ذات الصلة خارج الأمم المتحدة، بما في ذلك عملية قمم تسخير الذكاء الاصطناعي المسؤول في المجال العسكري، والإعلان السياسي بشأن الاستخدام العسكري المسؤول للذكاء الاصطناعي وخاصية التشغيل الذاتي، وفريق الخبراء الحكوميين المعني بالتكنولوجيات الناشئة في مجال منظومات الأسلحة الفتاكة الذاتية التشغيل. وترى جمهورية كوريا أن هذه المبادرات يعزز بعضها بعضًا ويكمل بعضها بعضًا.

إن حوكمة البيانات أمر بالغ الأهمية أيضاً. ونظرًا لأن البيانات تؤدي دورًا محوريًا في التدريب على منظومات الذكاء الاصطناعي ونشرها وتقييمها، يجب على أصحاب المصلحة المعنيين إجراء مزيد من النقاشات بشأن الآليات المناسبة لحوكمة البيانات، بما في ذلك وضع سياسات وإجراءات واضحة لجمع البيانات وتخزينها ومعالجتها وتبادلها وحذفها، فضلاً عن حماية البيانات.

الاتحاد الروسي

[الأصل: بالروسية]
[10 نيسان/أبريل 2025]

يرحب الاتحاد الروسي باتخاذ قرار الجمعية العامة 239/79 المؤرخ 24 كانون الأول/ديسمبر 2024، ويشرفه، وفقاً للفقرة 7 منه، أن يقدم إسهامه الوطني في تقرير الأمين العام إلى الجمعية العامة في دورتها الثمانين لكي تواصل الدول الأعضاء مناقشته.

مقدمة

يولي الاتحاد الروسي المسائل المتعلقة بتطبيق الذكاء الاصطناعي في المجال العسكري أهمية كبيرة. ونحن مهتمون بإجراء مزيد من المناقشات الموضوعية لهذه المسألة في المنتديات الدولية المتخصصة.

إننا نعتبر أن فريق الخبراء الحكوميين المعني بمنظومات الأسلحة الفتاكة الذاتية التشغيل، الذي أنشأته الأطراف المتعاقدة السامية في الاتفاقية المتعلقة بأسلحة تقليدية معينة، هو المنتدى الأفضل لإجراء هذه المناقشة. ففريق الخبراء الحكوميين تحديدا هو المدعو إلى الحفاظ على توازن معقول بين الشواغل الإنسانية والمصالح الدفاعية المشروعة للدول في ما يتعلق بهذه الأسلحة، وإلى اتخاذ قرارات توافقية. إنَّ

نظر الفريق في التطبيقات العسكرية للذكاء الاصطناعي واسع النطاق، فهو لا يقتصر على مسألة منظومات الأسلحة الفتاكة الذاتية التشغيل بل يتطرق أيضا إلى عدد من الجوانب الهامة (بما فيها الجوانب القانونية والتقنية والعسكرية) المتعلقة باستخدام التكنولوجيا لأغراض عسكرية.

وننوه بمناقشة هذا الموضوع في سياق الأنظمة القائمة في مجال تحديد الأسلحة ونزع السلاح وعدم الانتشار. وتركز هذه الورقة على تحليل المخاطر والفرص المواكِبة للذكاء الاصطناعي من حيث وفاء الدول الأطراف بالتزاماتها بموجب الصكوك القانونية الدولية ذات الصلة.

ونرحب باستعداد الدول الأعضاء للشروع في مناقشة موضوع التطبيقات العسكرية للذكاء الاصطناعي في هيئة نزع السلاح وذلك في إطار المناقشة المتعلقة بالتكنولوجيات الناشئة في سياق الأمن الدولي. والغرض من تبادل الآراء هذا هو الاتفاق على توصيات بشأن جوانب الذكاء الاصطناعي "العسكري" التي لم يجرَ تناولها في منتديات أخرى.

وفي سياق العمل في المنتديات الدولية المذكورة أعلاه، ينبغي إيلاء اهتمام خاص لوضع مصطلحات مشتركة، ولتطبيق القانون الدولي القائم، وللرقابة البشرية، والمساءلة، ولما تحمله التكنولوجيا من مخاطر وفرص.

تعريف

لا ينص القانون الدولي الموجود على تعريف متفق عليه لمنظومات الأسلحة والأعتدة العسكرية القائمة على الذكاء الاصطناعي، ما يجعل من الصعب تناوُل هذه المسألة. وعليه، فإن التوصل إلى فهم عملي مشترك لهذه الأدوات، وكذلك للمصطلحات المرتبطة بتطبيق هذه التكنولوجيا لأغراض عسكرية بشكل عام، سيكوّن صورةً أوضح عن الموضوع وعن آفاق مناقشته.

وينبغي للتعريف العملي أن يشمل ما يلي:

(أ) أن يتضمن توصيفا لأنواع منظومات الأسلحة والأعتدة العسكرية القائمة على الذكاء الاصطناعي والسمات الرئيسية المحدَّدة لاستخدامها؛

(ب) ألا يقتصر على الفهم الحالي لهذه الأدوات، بل أن يبين كيفية إمكان تطويرها في المستقبل؛

(ج) أن يقدم تعريفا مفهوما على الصعيد العالمي تضعه أوساط الخبراء بمن فيهم علماء ومهندسون وتقنيون وعسكريون ومحامون وأخصائيون في مجال الأخلاقيات؛

(د) ألا يُفهم منه بأنه يحد من التقدم التكنولوجي أو أنه يقوض البحوث في مجال التطبيقات السلمية لعلم الروبوتات والذكاء الاصطناعي؛

(هـ) ألا يعرِّف منظومات الأسلحة والأعتدة العسكرية القائمة على الذكاء الاصطناعي عن طريق وصف وظائفها حصراً.

كما ينبغي تجنُّب تصنيف هذه الأدوات بـ "سيئة" أو "جيدة"؛ وبتعبير آخر، لا ينبغي تصنيفها على أساس التفضيلات السياسية لمجموعة معينة من الدول.

ولا ينبغي تصنيف المنظومات العسكرية الحالية ذات الدرجة العالية من التشغيل الذاتي ضمن فئة "خاصة" تستلزم وضع قيود ومحظورات عاجلة. فهذا المستوى تحديدا من التشغيل الذاتي هو الذي يمكّن هذه المنظومات من العمل بفعالية في الحالات القتالية الدينامية وفي البيئات المختلفة، بما يضمن درجة مناسبة من التحديد والدقة، والامتثال تاليا لمبادئ وقواعد القانون الدولي، بما في ذلك القانون الدولي الإنساني.

منظومات الأسلحة والأعتدة العسكرية القائمة على الذكاء الاصطناعي في سياق القانون الدولي

من المسلَّم به عموما أن القانون الدولي القائم، بما في ذلك القانون الدولي الإنساني، يسري بالكامل على منظومات الأسلحة القائمة على الذكاء الاصطناعي.

إن الاتحاد الروسي يعتقد أن ما من أسباب مقنعة حاليا تدعو إلى فرض أي قيود أو محظورات جديدة على منظومات الأسلحة القائمة على الذكاء الاصطناعي، أو إلى تحديث أو تكييف القانون الدولي، بما في ذلك القانون الدولي الإنساني، في ما يتصل بهذه الوسائل. كما أن إجراء مناقشات تهدف إلى الاتفاق على نوع من "قواعد السلوك" أو المعايير والمبادئ الخاصة بالاستخدام "المسؤول" لمنظومات الأسلحة والأعتدة العسكرية القائمة على الذكاء الاصطناعي هو أمر سابق لأوانه. فمفهوم الاستخدام "المسؤول" للذكاء الاصطناعي الذي تروج له البلدان الغربية يستند إلى معايير مثيرة للجدل إلى حد كبير وغريبة عن القانون الدولي (بما في ذلك القانون الدولي الإنساني)، كما أنه يطرح العديد من التساؤلات ولا يحظى بتأييد توافقي من المجتمع الدولي.

ولا يمكن الاستناد إلى مبادئ الإنسانية، وما يمليه الضمير العام، وعنصر حقوق الإنسان باعتبارها الشرط الكافي المطلق والحصري لفرض قيود ومحظورات على أنواع معينة من الأسلحة والأعتدة العسكرية. وينبغي التصدي للشواغل المتعلقة بمنظومات الأسلحة والأعتدة العسكرية القائمة على الذكاء الاصطناعي من خلال تطبيق القواعد القانونية الدولية القائمة بحسن نية.

ويظل الامتثال الصارم لقواعد ومبادئ القانون الدولي، بما في ذلك القانون الدولي الإنساني، في حالات النزاع المسلح من أولويات الاتحاد الروسي. فالقوات المسلحة للاتحاد الروسي تتقيد تقيدا صارما بمعايير القانون الدولي الإنساني المنصوص عليها في الصكوك القانونية للاتحاد والمقاطعات. وتنعكس القضايا المتصلة بالامتثال للقانون الدولي الإنساني، بما فيها تلك المرتبطة باستخدام أنواع جديدة من الأسلحة، في القواعد التنظيمية والبرامج التدريبية لجميع فئات الأفراد العسكريين. وفي عام 2022 اعتُمدت ورقة مفاهيمية للقوات المسلحة للاتحاد الروسي حول تطوير واستخدام منظومات الأسلحة القائمة على الذكاء الاصطناعي.

ويراعي القانون الروسي بالكامل المبادئَ التوجيهية المتعلقة بأنظمة الأسلحة القائمة على الذكاء الاصطناعي التي توافقت عليها الدول الأطراف في الاتفاقية المتعلقة بأسلحة تقليدية معينة في عام 2019. ونحن نرى في مواصلة تبادل المعلومات حول التدابير العملية المحددة لتنفيذ هذه المبادئ التوجيهية على الصعيد الوطني وسيلةً لبناء الثقة وتعزيز الشفافية.

التحكّم في منظومات الأسلحة والأعتدة العسكرية القائمة على الذكاء الاصطناعي

إننا نعتبر أن ضرورة تحكّم البشر في تشغيل منظومات الأسلحة والأعتدة العسكرية القائمة على الذكاء الاصطناعي تشكل تقييدا كبيرا. لذا ينبغي لحلقة التحكم في هذه الأدوات أن تسمح بتدخل مشغِّل

بشري أو نظام تحكُّم عالي المستوى لتغيير طريقة تشغيل هذه المنظومات، بما في ذلك تعطيلها بشكل جزئي أو كامل.

ويرى الاتحاد الروسي أن البشر هم دائماً مسؤولون عن قرارات استخدام القوة. كما أن التحكّم الذي يمارَس يستند إلى كل المعلومات المتاحة عند اتخاذ القرار. وفي الوقت نفسه، ينبغي أن تُترك مسألة البت في الأشكال والأساليب المحددة للتحكم البشري لتقدير الدول، وينبغي ألا تتقرر بالضرورة من خلال الإدارة المباشرة وحدها.

ويمكن ممارسة التحكم بهذه المنظومات والأعتدة من خلال:

(أ) تحسين موثوقيتها وقدرتها على مواصلة العمل رغم الخلل؛

(ب) تقييد أنواع الأهداف؛

(ج) تقييد المدة والتغطية الجغرافية ونطاق التطبيق؛

(د) المسارعة إلى التدخل فيها وتعطيلها؛

(هـ) اختبار منظومات الأسلحة الفتاكة الذاتية التشغيل في ظل ظروف عملياتية تحاكي الواقع؛

(و) السماح للأشخاص الذين أتقنوا بنجاح إجراءات استخدام الأدوات القائمة على الذكاء الاصطناعي بتشغيلها (التحكّم فيها)؛

(ز) مراقبة تصنيع كلٍ من المكوّنات والجهاز ككل؛

(ح) مراقبة تفكيكِ كلٍ من المكوّنات والجهاز ككل والتخلصِ منهما.

إننا نعتبر أن من غير المناسب إقحام مفاهيم "التحكّم البشري الفعلي" و "أشكال ودرجات المشاركة البشرية" و "التحكّم والتقييم البشريين المناسبين للسياق" و "قابلية التنبؤ وقابلية الاعتماد، وقابلية التتبّع، وقابلية التفسير"، التي تروج لها دول معينة، في المناقشة لأن هذه المفاهيم لا مفاعيل قانونية لها عموما ولا تؤدي إلا إلى تسييس المناقشات.

المسؤولية

يعتقد الاتحاد الروسي أن الدول والأفراد (بما في ذلك الجهات المطوِّرة والجهات المنتجة) يتحملون في كل وقت بموجب القانون الدولي المسؤولية في عن قراراتهم المتعلقة بتطوير واستخدام منظومات الأسلحة والأعتدة العسكرية القائمة على الذكاء الاصطناعي. وتقع المسؤولية عن استخدام هذه الأدوات على عاتق المسؤول الذي يكلّفها بالمهمة ويصدر الأوامر باستخدامها. ولدى استخدام منظومات الأسلحة والأعتدة العسكرية القائمة على الذكاء الاصطناعي ينبغي لذلك المسؤول أن يمتلك المعرفة والمهارات المطلوبة في ما يتعلق بطريقة عملها وتشغيلها، وينبغي له أن يكون مسؤولا عن اتخاذ القرار بشأن مدى ملاءمة استخدامها والتخطيط لأشكال وأساليب استخدامها.

فرص وقيود منظومات الأسلحة والأعتدة العسكرية القائمة على الذكاء الاصطناعي

من الحقائق المعروفة جيدا أن منظومات الأسلحة والأعتدة العسكرية القائمة على الذكاء الاصطناعي يمكن أن تكون أكثر فعالية من المشغّل البشري في تنفيذ المهام ويمكن أن تقلل احتمال حصول أخطاء. وعلى وجه التحديد، إن هذه الأدوات قادرة على أن تحدّ بدرجة كبيرة من الآثار السلبية – في سياق القانون الدولي، بما في ذلك القانون الدولي الإنساني – المرتبطة بأخطاء ناجمة عن المشغّل، أو عن حالته العقلية والبدنية، أو عن معنوياته ومعتقداته الدينية والأخلاقية. ويمكن أن يضمن استخدامُ هذه الأدوات تحسينَ دقة استهداف الأسلحة للمرافق العسكرية، وأن يساعد في الحد من احتمال شن هجمات غير مقصودة على المدنيين والأعيان المدنية.

وينبغي لتقييم المخاطر المحتملة المتصلة باستخدام منظومات الأسلحة والأعتدة العسكرية القائمة على الذكاء الاصطناعي وبتدابير التخفيف من حدتها أن يكون جزءا من عمليات تصميم التكنولوجيات الجديدة وتطويرها واختبارها ونشرها في أي نوع من أنواع منظومات الأسلحة.

ويمكن التقليل إلى أدنى حد من المخاطر المتصلة بهذه الأدوات من خلال ما يلي:

(أ) ضمان الإدارة الفعالة لدورة صلاحيتها؛

(ب) إجراء اختبارات شاملة في جميع مراحل دورة صلاحيتها، بما في ذلك في بيئات تحاكي الواقع؛

(ج) تحسين موثوقيتها وقدرتها على مواصلة العمل رغم الخلل؛

(د) تحديد معايير الجاهزية؛

(هـ) توفير الحماية القصوى من الوصول غير المأذون إليها؛

(و) تدريب المشغّلين؛

(ز) إيلاء الأولوية لاستخدام تكنولوجيات الذكاء الاصطناعي في جمع ومعاملة المعلومات اللازمة دعماً لاتخاذ القرار العسكري؛

(ح) تيسير الرصد المستمر لعمليات هذه المنظومات من قِبل المشغّل وتمكين تحكّمه في إنهاء مهمة قتالية بأمر من المشغِّل؛

(ط) منع الجهات غير التابعة للدول من استعمال هذه الأدوات لأغراض غير مشروعة.

ويمكن اتخاذ هذه التدابير في كل مراحل دورة صلاحية (التطوير والإنتاج والتشغيل والتخلص) الأسلحة والأعتدة العسكرية والخاصة.

الخطوات المقبلة

نعتقد أنه سيكون من المفيد أن تواصل الدول النظر في القضايا المتعلقة بتطبيق الذكاء الاصطناعي للأغراض العسكرية ضمن فريق الخبراء الحكومي المعني بمنظومات الأسلحة الفتاكة الذاتية التشغيل، باعتباره المنتدى الدولي الأفضل لإجراء هذه المناقشة في سياق الأنظمة القائمة لتحديد الأسلحة

ونزع السلاح وعدم الانتشار، وكذلك في هيئة نزع السلاح. وفي الوقت نفسه، ينبغي للمناقشة في أحد المنتديات ألا تكرر تبادل الآراء الجاري بالفعل في منتديات موازية.

إننا نعارض تجزئة الجهود في هذا المجال. ويبدو أنه من غير المجدي إحالة قضية استخدام الذكاء الاصطناعي للأغراض العسكرية إلى أي منصات دولية أخرى، أو إنشاء منتديات إضافية للنظر فيها أو مناقشتها في منتدى ضيق دون مشاركة الأغلبية الساحقة من الدول الأعضاء في الأمم المتحدة (بما في ذلك الجهات المطوِّرة الرئيسية لمنظومات الأسلحة القائمة على الذكاء الاصطناعي، ومن ضمنها الاتحاد الروسي).

وبشكل خاص، إن المناقشات الدائرة حول هذا الموضوع في سياق "مؤتمرات القمة (غير الشاملة للجميع) حول الاستخدام المسؤول للذكاء الاصطناعي للأغراض العسكرية" التي نظمتها مجموعة من الدول الغربية، ومؤتمرات القمة حول الذكاء الاصطناعي بشكل عام، ليست بنّاءة. فهذه المناسبات ووثائقها الختامية لا تراعي آراء جميع الجهات المعنية ولا يمكن اعتبارها بالتالي أساساً يمكن البناء عليه لمواصلة عمل من شأنه أن يعكس فهماً مشتركاً للموضوع. وهي خلافية ولا تساعد على تضافر الجهود في هذا المجال.

وستكون لمحاولات "توحيد" النهُج الانفرادية المتبعة إزاء هذه القضايا في محافل بديلة، بما فيها "مؤتمرات القمة" هذه، التي تتجاوز المحافل المتخصصة المتعددة الأطراف، عواقب سلبية للغاية. فهي تنطوي على إمكانية التقويض الجدّي للعمل البنّاء والشامل للجميع المضطلع به في مجال الذكاء الاصطناعي "العسكري" وعلى تجزئة الجهود الرامية إلى إيجاد تفاهمات مشتركة ووضع توصيات في هذا المجال.

وفي سياق المناقشة في المنتديات الدولية المذكورة أعلاه، نعتقد أن من الضروري التركيز بشكل رئيسي على الاتفاق على مصطلحات متخصصة ونهُج مشتركة في ما يتعلق بتطبيق القانون الدولي القائم، بما في ذلك القانون الدولي الإنساني، على منظومات الأسلحة والأعتدة العسكرية القائمة على الذكاء الاصطناعي، والتركيز أيضا على ضمان التحكّم البشري بهذه الأدوات، وعلى المخاطر والفرص التي تحملها معها هذه التكنولوجيا.

إن الاتحاد الروسي يطلب إلى الأمين العام أن يأخذ في الاعتبار المقترحات أعلاه في تقريره الموضوعي المقدَّم عملا بالفقرة 7 من قرار الجمعية العامة 239/79، وأن يدرج هذه الوثيقة في مرفق ذلك التقرير.

صربيا

[الأصل: بالإنكليزية]
[4 نيسان/أبريل 2025]

يشكل تطوير الذكاء الاصطناعي وتطبيقه عاملاً مهماً من عوامل التغيير التي تطال طريقة إدارة العمليات العسكرية في عالم اليوم. فهو يوفر إمكانيات جديدة، ويطرح في الوقت نفسه تحديات جديدة أمام الاستقرار والسلام والأمن الدوليين في المجال العسكري. ولذلك، من الضروري الشروع في إنشاء إطار دولي مناسب لتنظيم تطبيقه.

1 - إمكانات ومزايا استخدام الذكاء الاصطناعي في المجال العسكري

قد يؤدي تطبيق الذكاء الاصطناعي في السياق العسكري غير الفتاك إلى تحسين العديد من مجالات العمليات العسكرية:

(أ) رفع مستوى الإلمام بحالة العمليات؛

(ب) تحسين عملية اتخاذ القرار من حيث الجودة والسرعة؛

(ج) الارتقاء بجودة البيانات الاستخباراتية والاستطلاع من خلال المعالجة السريعة للبيانات وإتاحة الكشف السريع عن التهديدات؛

(د) دعم حماية المدنيين وغير المحاربين في النزاعات العسكرية؛

(هـ) دعم عمليات وبعثات السلام من خلال رصد عمليات وقف إطلاق النار والتنبؤ بديناميات النزاع؛

(و) تحسين عمليات وإجراءات الصيانة التنبؤية وتحسين الخدمات اللوجستية من خلال تقليل التكاليف وتوفير الموارد.

2 - التحديات والتهديدات الرئيسية الناجمة عن استخدام الذكاء الاصطناعي في المجال العسكري

يشكل تطوير الذكاء الاصطناعي وإدماجه في المنظومات القتالية وغير القتالية تحديًا جسيما للسلام والاستقرار الدوليين، وكذلك للقانون الدولي الإنساني، وذلك في المجالات التالية أساسا:

(أ) المخاطر التقنية والأعطال الوظيفية الناجمة عن أخطاء التطبيق في بيئة دينامية، وهي مخاطر يمكن أن تهدد حياة البشر وتسبب أضرارا مادية وتؤثر على تنفيذ القانون الدولي الإنساني؛

(ب) المخاطر القانونية والأخلاقية المتعلقة بالامتثال للقانون الدولي، لا سيما فيما يتعلق بتنفيذ مبادئه، مثل التمييز والتناسب واتخاذ تدابير احترازية لدى الاستهداف؛

(ج) عدم وجود قواعد صريحة تحدد من يتحمل المسؤولية عن الأفعال والأنشطة التي يقوم بها الذكاء الاصطناعي؛

(د) قد تؤدي أوجه القصور في الخوارزميات، على الأرجح، إلى التحيز وارتكاب أخطاء في عملية اتخاذ القرار وإلى التمييز، لأن تطبيق مجموعات بيانات غير تمثيلية قد يؤدي إلى تحديد خاطئ للمدنيين أو قد يشكل تهديدا لمجموعات عرقية أو قومية؛

(هـ) قد يؤدي تطبيق خوارزميات الذكاء الاصطناعي إلى إيجاد انطباعات خاطئة لدى الأفراد المشاركين في عملية تنفيذ العمليات بأن مسؤوليتهم عنها مخففة؛

(و) المخاطر الاستراتيجية التي ينطوي عليها اتخاذ القرارات عن طريق الذكاء الاصطناعي، استناداً إلى فرضيات خاطئة؛

(ز) التقارب والتكامل مع التكنولوجيات الجديدة بصورة غير انتقائية، لا سيما في مجالات المعلومات والعمليات السيبرانية أو استخدام الوسائل النووية والكيميائية والبيولوجية؛

(ح) عدم وجود موظفين محترفين يقومون بالتطوير والتنظيم والتطبيق المسؤول لمنظومات الذكاء الاصطناعي في النزاعات؛

(ط) إساءة استخدام الذكاء الاصطناعي في العمليات الإعلامية عن طريق اختلاق معلومات مضللة ونشرها، مما قد يؤدي إلى إثارة النزاعات وتفاقم حدة التوترات.

3 - إنشاء إطار عمل قانوني وأخلاقي

بالنظر إلى المخاطر والتحديات التي جرى تقييمها، من الضروري وضع أطر قانونية وأخلاقية إلزامية داخل المجتمع الدولي من أجل القيام بما يلي:

(أ) تشجيع إجراء حوار داخل الأمم المتحدة والعمل على بدئه، بهدف زيادة الامتثال لمعايير القانون الدولي الإنساني، بما في ذلك وضع معايير وقواعد ومبادئ قانونية دولية تضمن أن يكون تطوير وتطبيق منظومات الذكاء الاصطناعي متوافقاً مع مبادئ القانون الدولي الإنساني (التمييز والتناسب واتخاذ التدابير الاحترازية لحماية الأفراد غير المشاركين في النزاعات المسلحة)؛

(ب) بدء عملية تقييم لشرعية تطبيق المنظومات والأسلحة على ضوء التطبيقات المقبولة المعتمدة للذكاء الاصطناعي؛

(ج) ضمان حماية حياة الأفراد وحريتهم أثناء النزاعات المسلحة وخصوصيتهم في وقت السلم، لا سيما في سياق المراقبة؛

(د) تعزيز آليات الأمم المتحدة من خلال بدء فرض النظر الإلزامي في مخاطر تطبيق الذكاء الاصطناعي للأغراض العسكرية، ورفع مستوى مؤتمر نزع السلاح، ومواءمة عمل هيئة نزع السلاح، وإنشاء هيئات متخصصة جديدة تابعة للأمم المتحدة، وتوسيع نطاق مبادرات الأمم المتحدة الحالية للاستخدام المسؤول للذكاء الاصطناعي؛

(هـ) بدء حوار في الأمم المتحدة لتحديد الاستخدام المسؤول للذكاء الاصطناعي في المجال العسكري ووضع بروتوكولات أمنية لتطبيقه (الاختبار والتقييم والاعتماد والتحقق)؛

(و) وضع تدابير لمواءمة إشراك القطاع الخاص مع مبادئ القانون الدولي الإنساني أثناء تطوير وإنشاء وتطبيق منظومات وخدمات الذكاء الاصطناعي في المجالات العسكرية؛

(ز) توسيع نطاق معاهد الأمم المتحدة ووثائقها الحالية بشأن التوصيات المتعلقة بأخلاقيات تطوير وتطبيق الذكاء الاصطناعي لتشمل مبادئ توجيهية محددة بشأن السلوك في النزاعات.

ويستدعي تطبيق منظومات الذكاء الاصطناعي في سياق النزاعات المسلحة الدولية من المجتمع الدولي اتخاذ إجراءات واسعة متعددة الأطراف، بهدف تعزيز المسؤولية عن استخدامها. وينبغي للأمم المتحدة أن تضطلع بدور قيادي في تحفيز إجراء حوار ووضع المعايير وبناء قدرات المجتمع الدولي من أجل منع التجزئة وتحقيق الإدارة السليمة.

سنغافورة

[الأصل: بالإنكليزية]
[11 نيسان/أبريل 2025]

ما فتئت سنغافورة، باعتبارها دولة صغيرة، تؤيد النظامَ المتعدد الأطراف القائم على القواعد ودورَ الأمم المتحدة. فالأمم المتحدة توفر الأساس للقانون الدولي والقواعد الدولية. والمؤسسات والنظم والقوانين المتعددة الأطراف ضرورية لبقاء جميع الدول، ولا سيما الدول الصغيرة.

وتعتقد سنغافورة أن قدرات الذكاء الاصطناعي في المجال العسكري، بما في ذلك المنظومات القائمة على الذكاء الاصطناعي، يجب أن تُطبق بطريقة مسؤولة طوال دورة حياتها بالكامل وبما يمتثل للقانون الدولي الساري، ولا سيما القانون الدولي الإنساني.

ويمكن للذكاء الاصطناعي أن يجلب فوائد في المجال العسكري فيما يتعلق بتعزيز الدقة والإلمام بالحالة، وبالتالي الحد من الأضرار الجانبية التي تلحق بالمدنيين و/أو الأعيان المدنية. ومع ذلك، يمكن للذكاء الاصطناعي أيضًا أن تسبب مخاطر تصعيد النزاع وسوء التقدير في ظل غياب أطر حوكمة مناسبة. وفي هذا الصدد، تعتقد سنغافورة أن من المهم أن يتحاور المجتمع الدولي بشأن هذا الموضوع.

النهج الذي تتبعه سنغافورة إزاء حوكمة الذكاء الاصطناعي في المجال العسكري

أحد الأهداف الرئيسية لاستراتيجية سنغافورة الوطنية للذكاء الاصطناعي 2,0 هو تشجيع تهيئة بيئة تسودها الثقة تحمي المستخدمين وتيسر الابتكار. وتحقيقاً لهذه الغاية، تعمل مختلف القطاعات الحكومية، بما في ذلك الدفاع، على وضع أطر عمل لحوكمة الذكاء الاصطناعي لكي يتسنى تسخير فوائد الذكاء الاصطناعي، مع ضمان تخفيف الأضرار المحتملة لاستخدامه.

وبفضل المشاورات مع خبراء تكنولوجيا الدفاع والمخططين العسكريين وخبراء القانون الدولي والمتخصصين في مجال السياسات، وضعت سنغافورة مبادئ وطنية متعلقة بالذكاء الاصطناعي في المجال العسكري، أُعلنت عنها في عام 2021، وهي تتناول أربعة مجالات اهتمام رئيسية:

(أ) **المسؤولية** – أولا، يجب معالجة مخاطر ظهور سلوك ناشئ للذكاء الاصطناعي. ويجب أن يكون لمنظومات الذكاء الاصطناعي استخدامات مقصودة محددة جيدا، وأن يتحمل كل من المطوِّرين والمستخدمين مسؤولية النتائج المترتبة على هذه المنظومات.

(ب) **الموثوقية** – ثانيا، لابد من التصدي للمخاطر المتعلقة بحدوث أخطاء في مخرجات منظومات الذكاء الاصطناعي أو بعدم اتسامها بالدقة. وينبغي اختبار هذه المنظومات والتأكد من أنها على المستوى المناسب لاستخداماتها المقصودة. وينبغي أن تكون مصمَّمة بحيث يكون تحيزها غير المقصود عند أدنى مستوى وبحيث تنتج مخرجات متسقة.

(ج) **المتانة** – ثالثا، يجب التصدي للمخاطر الناجمة عن استغلال الذكاء الاصطناعي من قِبَل جهات فاعلة ذات نوايا خبيثة. وينبغي تصميم منظومات الذكاء الاصطناعي مع أخذ التهديدات السيبرانية وتهديدات الذكاء الاصطناعي ذات الطابع العدائي في الحسبان. وبغية التصدي للطابع المبهم

لهذه المنظومات الذي تشبه "الصندوق الأسود"، ينبغي أن تكون عملية تطويرها موثقة جيدا لدعم قابليتها للتفسير.

(د) **الأمان** – رابعا، يجب التركيز على خطر تَعطُّل الذكاء الاصطناعي في السياقات البالغة الأهمية للسلامة. وينبغي أن تكون منظومات الذكاء الاصطناعي آمنة للاستخدام، ليس من حيث المنصات المنشورة فحسب، بل أيضا من حيث ما يحيط بها من أصول وأفراد.

وقد استرشدت سنغافورة بهذه المبادئ التوجيهية في نهج الحوكمة الذي تتبعه سنغافورة في تطوير واختبار وتدريب ونشر المنظومات القائمة على الذكاء الاصطناعي للأغراض العسكرية.

المبادرات الدولية والإقليمية بشأن الذكاء الاصطناعي في المجال العسكري

شاركت سنغافورة بنشاط في المبادرات الدولية بشأن حوكمة الذكاء الاصطناعي في المجال العسكري. ففي عام 2023، أيدت سنغافورة النداء للعمل من أجل الذكاء الاصطناعي المسؤول في المجال العسكري، والإعلان السياسي بشأن الاستخدام العسكري المسؤول للذكاء الاصطناعي وخاصية التشغيل الذاتي. وفي عام 2024، شاركت سنغافورة في استضافة قمة تسخير الذكاء الاصطناعي المسؤول في المجال العسكري في سيول، جمهورية كوريا، حيث أيدنا مخطط العمل من أجل الذكاء الاصطناعي المسؤول في المجال العسكري.

وتدرك سنغافورة أيضًا أهمية المبادرات الإقليمية لضمان إجراء مناقشات شاملة للجميع ومحددة السياق بشأن الذكاء الاصطناعي في المجال العسكري. وقد شاركت سنغافورة في استضافة المشاورات الإقليمية لعام 2024 بشأن الذكاء الاصطناعي المسؤول في المجال العسكري في آسيا، وهي مشاورات وفرت محفلا لبلدان المنطقة لتبادل وجهات النظر، بما في ذلك بشأن الفرص والمخاطر التي ينطوي عليها الذكاء الاصطناعي في المجال العسكري.

وفي شباط/فبراير 2025، اعتمدت سنغافورة وغيرها من الدول الأعضاء في رابطة أمم جنوب شرق آسيا بياناً مشتركاً بشأن التعاون في مجال الذكاء الاصطناعي في قطاع الدفاع في معتكف اجتماع وزراء دفاع الرابطة في بينانغ بماليزيا. وفي البيان، تعهد وزراء دفاع الرابطة بتعزيز استخدام الذكاء الاصطناعي على نحو مسؤول وخاضع للمساءلة، وتعميق الفهم والوعي على الصعيد الإقليمي بآثار الذكاء الاصطناعي في قطاع الدفاع من خلال تبادل المعلومات، وتبادل أفضل الممارسات والدروس المستفادة فيما بين الدول الأعضاء في الرابطة.

سبل المضي قدما في المناقشات بشأن الذكاء الاصطناعي والسلام والأمن الدوليين في الأمم المتحدة

تعتقد سنغافورة أن أي مناقشات أخرى تُجرى للبناء على دعم المجتمع الدولي لهذا القرار ينبغي أن تكون ذات طابع مفتوح وشامل للجميع. وفي هذا الصدد، فإننا نؤيد إنشاء فريق عامل مفتوح باب العضوية في إطار الأمم المتحدة يركز على الذكاء الاصطناعي في المجال العسكري. وإذا أنشئ هذا الفريق العامل المفتوح باب العضوية، ينبغي أن يتبنى نهجا متعدد أصحاب المصلحة يشارك فيه، من بين آخرين، خبراء التكنولوجيا والمخططون العسكريون والخبراء القانونيون الدوليون والمتخصصون في مجال السياسات. ونؤكد من جديد التزامنا بالعمل مع جميع الدول الأعضاء للنهوض بالتطبيق المسؤول للذكاء الاصطناعي في المجال العسكري.

إسبانيا

[الأصل: بالإسبانية]
[11 نيسان/أبريل 2025]

مقدمة

ينطوي الذكاء الاصـــطناعي على ثورة في جميع المجالات، بما في ذلك الأمن والدفاع. ويحمل تطويره واستخدامه تقدما كبيرا وفرصا عظيمة، ويطرح في الوقت نفسه العديد من التحديات.

واعتماد هذه التكنولوجيا من قبل القوات المســلحة لا يعيد تعريف الطريقة التي تجري بها العمليات العسكرية فحسب، بل إنه يحدث أيضاً تحولا في التوازن الاستراتيجي العالمي.

ويســتند تطوير وزارة الدفاع للذكاء الاصــطناعي وإدماجه إلى الاســتخدام العســكري المســؤول والأخلاقي والشرعي الذي يتماشى مع القانون الدولي الإنساني ويضمن احترام حقوق الإنسان.

والذكاء الاصـــطناعي يغير المفهوم التقليدي للقوة العســـكرية والأمن، ويوفر قدرات متقدمة لجمع البيانات وتحليلها واتخاذ القرارات وتنفيذ العمليات في بيئات متعددة المجالات. وينطوي ذلك على نقلة نوعية في الطريقة التي تتعامل بها الدول مع الدفاع والأمن، وييسر تنظيم استجابات أسرع وأدق للتهديدات الناشئة.

وفي المجال العسكري، يُحدث الذكاء الاصطناعي تأثيرا كاسحا في ساحات المعارك التي لا يمكن التنبؤ بها، مما يؤدي إلى نقلة نوعية في تخطيط العمليات العســكرية وتنفيذها. ويؤثر الذكاء الاصـــطناعي أيضًـــا على مجالات أخرى في المجال العســـكري: اللوجســـتيات والتدريب وإدارة المعلومات وتفســـيرها والاستخبارات والرصد وتحديد الهدف والاستطلاع.

وتجدر الإشارة إلى أن إسبانيا، تماشيًا مع التزامها بالاستخدام المسؤول للذكاء الاصطناعي، هي البلد المضــيف لقمة تســخير الذكاء الاصــطناعي المســؤول في المجال العســكري لعام 2025، وقد أيدت "الدعوة إلى العمل" (لاهاي، 2023) و "مخطط العمل" الذي قُدم في القمة الأخيرة في عام 2024.

الإطار المفاهيمي والتنظيمي لوزارة الدفاع

تسترشد وزارة الدفاع في تطوير الذكاء الاصطناعي ونشره وتطبيقه بمجموعة من المبادئ الأساسية التي تضـــمن اســـتخدامه الآمن والأخلاقي بما يتوافق مع اللوائح الوطنية والدولية. وتهدف هذه المبادئ، المنصوص عليها في "استراتيجية تطوير الذكاء الاصطناعي وإدماجه واستخدامه من قبل وزارة الدفاع" (التي وُضعت بموجب القرار 2023/11197 الصادر عن أمانة الدولة للدفاع) والتي تتماشى مع استراتيجية الذكاء الاصطناعي لمنظمة حلف شمال الأطلسي (ناتو) لعام 2021 (تمت مراجعتها في عام 2024)، إلى تعظيم الفرص التي يوفرها الذكاء الاصـــطناعي في مجال الدفاع مع الســـعي إلى التخفيف من المخاطر المرتبطة باستخدامه في المجال العسكري:

- الشـــرعية: يجب تطوير تطبيقات الذكاء الاصـــطناعي واســـتخدامها وفقًا للقانون الوطني والدولي المعمول بهما، بما في ذلك الإعلان العالمي لحقوق الإنسان والقانون الدولي الإنساني.
- المسؤولية والمساءلة البشرية: يجب إفساح المجال في أي تطوير للذكاء الاصطناعي، وكذلك في استخدامه، لرقابة بشرية واضحة من أجل ضمان المساءلة على النحو المناسب وإسناد المسؤولية.

- قابلية التفسير والتتبع: يجب أن تكون تطبيقات الذكاء الاصطناعي، بما في ذلك استخدام المنهجيات والمصادر والإجراءات القابلة للتدقيق، مفهومة وشفافة للموظفين المعنيين.
- الموثوقية والشفافية: ينبغي أن تكون تطبيقات الذكاء الاصطناعي مصممة لحالات استخدام دقيقة ومحدودة ومحددة بشكل جيد، وينبغي توفير المعلومات لتعزيز الفهم العام لهذه التطبيقات من قبل جميع أصحاب المصلحة. ويجب أن تخضع سلامة وأمن وكفاءة هذه القدرات للاختبار والضمانات في ظل حالات الاستخدام هذه طوال دورة حياتها.
- إمكانية التحكّم: يجب تطوير تطبيقات الذكاء الاصطناعي واستخدامها وفقًا لمهام التصميم المتوخاة، ويجب أن تتضمن القدرة على اكتشاف العواقب غير المقصودة وتجنبها. ويجب تمكين آليات تتيح الفصل أو التعطيل عند ملاحظة سلوك غير مقصود أو غير مرغوب فيه.
- تخفيف التحيز: يجب اتخاذ جميع التدابير اللازمة للتقليل إلى أدنى حد من الأخطاء والميول الذاتية في تطوير واستخدام الذكاء الاصطناعي.
- الخصوصية: يجب أن يحترم تطوير التطبيقات القائمة على الذكاء الاصطناعي وتنفيذها واستخدامها خصوصية الأفراد، بدءاً من تصميمها وطوال دورة حياتها.

وفيما يتعلق بالإطار التنظيمي، يجري العمل على وضع مجموعة من المعايير وأفضل الممارسات بشأن تطوير وتنفيذ واستخدام الذكاء الاصطناعي في المجال العسكري من قبل وزارة الدفاع لضمان الاستخدام المسؤول والفعال للذكاء الاصطناعي وفقًا للأطر القانونية الوطنية والدولية، وعلى وجه الخصوص، مع الامتثال الصارم للقانون الدولي الإنساني وحقوق الإنسان.

الفرص

تركز وزارة الدفاع تطويرها لقدرات الذكاء الاصطناعي على مجالات متنوعة لتحسين فعالية القوات المسلحة. ووفقًا للاستراتيجية، يركز استخدام الذكاء الاصطناعي على مجالات العمليات والاستخبارات والخدمات اللوجستية والأمن السيبراني ودعم اتخاذ القرار.

وسيساعد الذكاء الاصطناعي في زيادة دقة وسرعة وفعالية اتخاذ القرار أثناء العمليات العسكرية، مع مراعاة القانون الدولي الإنساني في جميع الأوقات، بهدف تنفيذ المهام بكفاءة أكبر وتقليل المخاطر التي تتعرض لها القوات، فضلاً عن المساعدة في تعزيز حماية المدنيين والأعيان المدنية في النزاعات المسلحة.

وحيث أن الذكاء الاصطناعي يتيح تحليل كميات كبيرة من البيانات في الوقت الفعلي، فهو يحسن القدرات المتعلقة بتقدير الحالة العسكرية والتصدي للتهديدات، مما يعزز أمن العمليات. وتتيح كل هذه التحسينات في القدرات التحكم البشري في جميع الأوقات ولا تفوض المسؤولية للآلات.

وفيما يتعلق بالتدريب والتعليم العسكريين، يجري العمل في إطار كليات الأركان العامة الأوروبية ("مجموعة القادة الخمسة") (إسبانيا وألمانيا وإيطاليا وبريطانيا العظمى وفرنسا)، على إنشاء فضاء تعاوني يتعلق بالذكاء الاصطناعي في التربية العسكرية.

وتتعاون إسبانيا أيضًا مع مجلس استعراض البيانات والذكاء الاصطناعي التابع لمنظمة حلف شمال الأطلسي بشأن الاستخدام المسؤول للبيانات والذكاء الاصطناعي في المجال العسكري.

وبالإضافة إلى ذلك، أعلنت وزارة الدفاع عن استثمارات استراتيجية في مناطق محددة بهدف تعزيز المشاريع المتعلقة بالذكاء الاصطناعي والتكنولوجيات المتقدمة الأخرى. ولا تسعى هذه الاستثمارات إلى تقوية هذا القطاع فحسب، بل كذلك إلى تعزيز التنشيط الصناعي في المناطق الإقليمية الجديدة.

التحديات

يجب أن يتماشى تطوير الذكاء الاصطناعي وتطبيقه في المجال العسكري مع الأطر التنظيمية الوطنية والدولية، بما في ذلك احترام تنفيذ القانون الدولي الإنساني، وبالتالي تعزيز الجهود الرامية إلى ضمان وجود تحكم بشري فعال في القرارات الحاسمة المرتبطة باستخدام الذكاء الاصطناعي في العمليات العسكرية.

وفيما يتعلق بالخصوصية وحماية البيانات، تطرح العمليات الضخمة لجمع البيانات وتجهيزها من أجل تدريب نماذج الذكاء الاصطناعي مخاطر فيما يتعلق بحماية البيانات الشخصية وأمن المعلومات.

السلامة والموثوقية

يتعلق التحدي الأكبر بالاستخدام الآمن والموثوق للذكاء الاصطناعي. وتتمثل المخاطر الرئيسية في هذا الصدد فيما يلي:

- يمكن أن تنطوي بيانات التدريب المتصلة بخوارزميات الذكاء الاصطناعي على أشكال من التحيز، مما قد يؤدي إلى قرارات خاطئة أو عواقب غير مقصودة.
- يمكن أن يؤدي التدريب الضعيف لنماذج الذكاء الاصطناعي إلى أخطاء في التفسير، مع عواقب كارثية محتملة على العمليات العسكرية.
- يمكن استهداف منظومات الذكاء الاصطناعي بهجمات سيبرانية يمكن أن تتلاعب بسلوكها أو تعطلها.
- هناك خطر تسميم البيانات، حيث تقوم جهات خبيثة بتغيير مجموعات بيانات التدريب للتسبب في فشل الخوارزمية.

وفي إسبانيا، يخضع تطوير الذكاء الاصطناعي في المجال العسكري لمبادئ المساءلة والمراقبة المستمرة، ويتم تنفيذ آليات تقييم المخاطر والتدقيق والتتبع في كل مرحلة من مراحل دورة حياة النظام. ويجب أن يسمح أي تطوير أو استخدام للذكاء الاصطناعي برقابة بشرية واضحة من أجل ضمان المساءلة المناسبة وإسناد المسؤولية، والسماح بإمكانية تتبع واضح للإجراءات البشرية المتعلقة بأنشطة الذكاء الاصطناعي والمتخذة بالتوازي معها، دون تفويض القرارات النهائية للآلات.

ويجب كذلك أن يكون الذكاء الاصطناعي موثوقاً وقابلاً للتنبؤ، مع الحفاظ على مستوى من التشغيل الذاتي يتم التحكم فيه والإشراف عليه من قبل مشغلين مدربين.

ويجب تقييم أي حل من حلول الذكاء الاصطناعي في وسط مختلف عن ذلك الذي تم تدريبه فيه، ويجب أن يخضع لاختبار غير وظيفي – اختبار الحمل والضغط والأداء في ظل سيناريوهات محددة ومتغيرة – لدراسة سلوكه والانحرافات المسموح بها.

وبالإضافة إلى ذلك، يجب أن تخضع قدرات الذكاء الاصطناعي هذه لاختبارات صارمة وعمليات تدقيق مستمرة طوال دورة حياتها، مما يتيح الكشف المبكر عن الأخطاء المحتملة وتحسين موثوقيتها التشغيلية. ويجب تنفيذ بروتوكولات الرقابة والتحكم البشريين في جميع مراحل النشر، بما يضمن عدم تفويض القرارات الحاسمة للذكاء الاصطناعي حصرياً. وفي هذا الصدد، تُبذل الجهود في الوقت الحالي لضمان مصادقة كيانات معترف بها على عمليات تطوير الذكاء الاصطناعي.

ومن أجل تحسين قوة المنظومات القائمة على الذكاء الاصطناعي وحمايتها من الإجراءات الخارجية، يعد إدماج الأمن منذ مرحلة التصميم أمرا أساسيا، فهو يضمن مقاومة هذه المنظومات للهجمات السيبرانية والتلاعب من قبل الخصوم ويكفل سلامة البيانات والنماذج المستخدمة.

ويمكن استهداف الذكاء الاصطناعي بهجمات مثل تلك التي تهدف إلى تسميم البيانات أو التلاعب بالنماذج؛ ولذلك يلزم إجراء مراقبة مستمرة لأداء النظام واختبارات التحقق من الصحة وعمليات التدقيق على نحو منتظم. وينبغي تعزيز عمليات وضع خطط النسخ الاحتياطي واستعادة القدرة على العمل بعد الكوارث، بما يضمن إمكانية تشغيل المنظومات في السيناريوهات غير المواتية.

وينبغي كذلك تعزيز التعاون مع وكالات الأمن السيبراني وخبراء الذكاء الاصطناعي لضمان أن يكون لدى القوات المسلحة أفضل الأدوات والاستراتيجيات لحماية هذه المنظومات من التهديدات الخارجية وضمان موثوقيتها التشغيلية.

وتعد المواهب وتدريب الأفراد على هذه التكنولوجيات أمرا بالغ الأهمية، ويشكل ذلك أحد المجالات الأربعة التي تركز عليها وزارة الدفاع، لضمان فهم المشغلين لنطاق وحدود هذه المنظومات وإمكانية تدخلهم في حالة حدوث انحرافات في سلوكها. إن تدريب الأفراد وتوعيتهم في مجال الاستخدام القانوني والأخلاقي للذكاء الاصطناعي أمر أساسي للتخفيف من المخاطر المرتبطة بالتحيز، وضمان أن يكون استخدام الذكاء الاصطناعي في القوات المسلحة موضوعيًا وموثوقًا وممتثلًا للوائح الوطنية والدولية، وخاصة القانون الدولي الإنساني.

ويجري حالياً إعداد دليل للممارسات الجيدة، يمكن أن يكون بمثابة أساس لوثيقة تتضمن مساهمات من جميع مجالات عمل وزارة الدفاع. وقد جرى نشر الممارسات الجيدة للاستخدام المسؤول للذكاء الاصطناعي في المجال العسكري التي اقترحتها منظمة حلف شمال الأطلسي. وأحد الأمثلة على ذلك هو تقييم الذكاء الاصطناعي المسؤول ومجموعة أدواته وفقا لمنظمة حلف شمال الأطلسي بهدف تفعيل مبادئ الاستخدام المسؤول للذكاء الاصطناعي التي اعتمدها حلف شمال الأطلسي، وتشمل المشروعية والمسؤولية وإمكانية التتبع والموثوقية وإمكانية التحكّم والتخفيف من التحيز.

سويسرا

[الأصل: بالإنكليزية]
[11 نيسان/أبريل 2025]

1 - الفرص والمخاطر

من المرجح أن يُحدث الذكاء الاصطناعي تحولاً في العديد من جوانب الشؤون العسكرية. فهو يبشر بدعم المهام والعمليات العسكرية، على سبيل المثال من خلال تعزيز الموثوقية والكفاءة والدقة والسلامة

والمتانة. ومن المجالات الرئيسية التي سيشملها ذلك الإلمام بالحالة، وصنع القرار، والاستخبارات، والمراقبة والاستطلاع، والخدمات اللوجستية وسلاسل الإمداد، والتدريب والمحاكاة، والقيادة والتحكم من خلال تحليل مجموعات البيانات الكبيرة وتمكين اتخاذ قرارات أسرع وأكثر استنارة. وعلى سبيل المثال، في مجال المراقبة والاستطلاع، يمكن للذكاء الاصطناعي أن يجري تحليلا للصور التي تلتقطها الطائرات المسيرة والسواتل بهدف اكتشاف التحركات بسرعة تفوق سرعة المحللين البشريين. ويمكن للذكاء الاصطناعي أيضا أن يساعد على التعرف على الأهداف من خلال معالجة بيانات أجهزة الاستشعار للتمييز بين القوات الصديقة والمعادية. وفي مجال الخدمات اللوجستية، يمكن للذكاء الاصطناعي تحسين سلاسل الإمداد إلى المستوى الأمثل والتنبؤ بأعطال المعدات وضمان وصول الموارد إلى المكان المناسب في الوقت المناسب. وفيما يتعلق بدعم اتخاذ القرار، يمكن لتمارين المحاكاة التي يتم القيام بها من خلال الذكاء الاصطناعي أن تزود القادة برؤى تنبؤية ونتائج محتملة يُسترشد بها في التخطيط الاستراتيجي. وتوفر نظم التدريب والمحاكاة التي تعمل بالذكاء الاصطناعي بيئات واقعية وقابلة للتكيف مع الظروف المحيطة بها تهيئ الجنود بشكل أفضل. وأخيراً، يمكن للذكاء الاصطناعي أن يدعم القيادة والتحكم من خلال ترشيد تدفق المعلومات وتحسين عملية صنع القرار وتعزيز التنسيق بين الوحدات. ويمكن للذكاء الاصطناعي أن يساعد في الكشف عن التهديدات، والأمن السيبراني، وحفظ السلام، والتحقق من تحديد الأسلحة، ووقف التصعيد في النزاعات من خلال نظم الإنذار المبكر، والتحليلات التنبؤية، وآليات الرصد، مما يساعد على تعزيز الاستقرار والأمن. ومع ذلك، وإن كانت هذه التطورات قد تعود بالفائدة على القوات المسلحة، فإن دمج الذكاء الاصطناعي في المجال العسكري يثير أيضًا عدة شواغل مهمة ومخاطر محتملة.

وعندما يُستخدم الذكاء الاصطناعي بمسؤولية في النزاعات المسلحة، قد يكون بمقدوره الإسهام في تعزيز الامتثال للقانون الإنساني الدولي وتعزيز حماية المدنيين والأعيان المدنية، على سبيل المثال من خلال تحسين تقييمات المخاطر أو زيادة دقة الاستهداف للحد من الأضرار الجانبية. ومع ذلك، فإن عدة أشكال من الذكاء الاصطناعي في المجال العسكري في النزاعات المسلحة، لا سيما تلك التي تنطوي على تطبيقات عالية الخطورة، تثير أيضًا شواغل خطيرة يجب معالجتها، تتنوع ما بين شواغل قانونية وإنسانية وأخلاقية وأمنية وشواغل متعلقة بالاستقرار الاستراتيجي، ومنها على سبيل المثال:

- **أخطاء اختيار الهدف** – قد يفلح الذكاء الاصطناعي، من الناحية التقنية، في تحديد الأهداف أو الأفراد على أساس بيانات التدريب، لكنه يواجه صعوبة خاصة في فهم السياق وفي إصدار أحكام قيمية، وهما أمران لازمان للامتثال للقانون الدولي، وهو ما قد يؤدي إلى تحديد خاطئ للأجسام أو الأشخاص واتخاذهم أهدافا عسكرية، وبالتالي إلى ضربات غير قانونية أو غير مقصودة.

- **مخاطر التصعيد** – في خضم أزمة سريعة التطور، قد توصي أداة دعم اتخاذ القرار الغامضة الشبيهة بالصندوق الأسود باتخاذ إجراءات عدوانية دون تقديم تعليل واضح. وفي ظل غياب قابلية التفسير، قد يتبع القادة إما توجيهات معيبة بشكل أعمى أو قد يضيعون وقتاً ثمينا في التشكيك فيها.

- **سوء تفسير النوايا** – قد تثير منظومة الذكاء الاصطناعي التي تقيّم المخاطر المرتبطة بتصرفات الأشخاص و/أو الأجسام شواغل (قانونية وأمنية)، لا سيما عندما تستند التقييمات إلى أنماط مستمدة من سلوكيات وسياقات السابقة دون تحكم وتقدير بشريين مناسبين للسياق. فعلى سبيل المثال، قد تخطئ منظومة الذكاء الاصطناعي التي تراقب سلوك الخصم في تصنيف التحركات

الروتينية للقوات على أنها تحركات عدائية، بسبب بيانات معيبة، مما قد يدفع إلى اتخاذ إجراءات استباقية وتصعيد غير مقصود.

وتبرز هذه المخاطر ضرورة الالتزام بضمان الامتثال للقانون الدولي القائم، لا سيما القانون الدولي الإنساني، وتبرز أيضًا الحاجة الملحة إلى إجراء مزيد من الحوار والدراسة لهذه المسألة لتحسين فهم المخاطر والتحديات والتدابير الضرورية الممكنة، وكذلك للنظر في ضرورة استحداث هياكل حوكمة معيارية إضافية وفي قيمته المضافة وجدواه. ويمكن أن يشمل ذلك التشريعات الوطنية، أو وضع أفضل الممارسات، أو القواعد أو المعايير أو الصكوك الدولية، أو وضع مبادئ توجيهية عملياتية.

2 - الإطار القانوني

لا يتم تطوير واستخدام الذكاء الاصطناعي، على غرار أي تكنولوجيا أخرى، في فراغ قانوني. ويجب أن يتم تطوير الذكاء الاصطناعي في المجال العسكري ونشره واستخدامه بالامتثال الكامل للقانون الدولي القائم، لا سيما ميثاق الأمم المتحدة والقانون الدولي الإنساني والقانون الدولي لحقوق الإنسان والأطر القانونية الأخرى ذات الصلة. ويجب ألا تطعن أي تكنولوجيا إطلاقا في صلاحية القانون الدولي. وينطبق القانون الدولي، ولا سيما ميثاق الأمم المتحدة برمته والقانون الدولي لحقوق الإنسان والقانون الدولي الإنساني، ويجب مراعاته والامتثال له.

ويجب على الدول وأطراف النزاع احترام وضمان احترام القانون الدولي الإنساني في جميع الظروف، بما في ذلك عند استخدام الذكاء الاصطناعي في العمليات العسكرية. وبالتالي، ينبغي تصميم الذكاء الاصطناعي في المجال العسكري بحيث يعزز الامتثال للقانون الدولي الإنساني وحماية المدنيين والأعيان المدنية. ويمكن تحقيق ذلك، على سبيل المثال، من خلال ضمان أن تعطي منظومات الذكاء الاصطناعي الأولوية للدقة وتقليل الأضرار والخضوع للمساءلة، بسبل منها إجراء عمليات صارمة لدى اختيار الأهداف والموافقة عليها والتحقق منها. وعلاوة على ذلك، ينبغي استخدام الذكاء الاصطناعي بطريقة تعزز تنفيذ الالتزام القانوني باتخاذ جميع الاحتياطات الممكنة في العمليات العسكرية، بما في ذلك تجنب الضرر العرضي أو على الأقل التقليل منه إلى أدنى حد، وذلك من خلال دعم القادة في حماية المدنيين والأعيان المدنية طوال سير الأعمال العدائية، على سبيل المثال، من خلال تحسين تقييمات المخاطر.

وأحد مجالات العمل الرئيسية هو ضمان أن يُصمم الذكاء الاصطناعي في المجال العسكري وأن يُدرب باستخدام مجموعات بيانات تمكّن من استخدامه بالامتثال الكامل للقانون الدولي. ولا يقتصر الأمر على سير الأعمال العدائية، إذ يجب أن يمتثل الذكاء الاصطناعي في المجال العسكري لجميع قواعد ومبادئ القانون الإنساني الدولي ذات الصلة، إذا ما استُخدم لأداء مهام أخرى يحكمها القانون الدولي الإنساني، على سبيل المثال فيما يتعلق باحتجاز واعتقال الأشخاص أو فيما يتعلق بتدابير فض الشغب وفرض الأمن العام في الأراضي المحتلة.

وعند تطوير الذكاء الاصطناعي واستخدامه في المجال العسكري، هناك خطر يتمثل في أن تصبح التفسيرات القانونية المفرطة في التساهل – مثل التعريف الفضفاض للأهداف المشروعة أو رفع عتبات الضرر العرضي المقبول – جزءًا لا يتجزأ من تصميم النظام أو بيانات التدريب. وإذا ما طُبقت هذه التفسيرات على نطاق واسع، فإنها قد تقوض تدريجياً الغرض الحمائي للقانون الدولي الإنساني وقد تزيد إلى حد كبير من الضرر الذي يلحق بالمدنيين. ويبرز هذا الخطر أهمية الحفاظ على سلامة المعايير القانونية،

التي يجب أن تظل، من الآن فصاعدا، اعتبارًا محوريًا في حوكمة الذكاء الاصطناعي وتصميمه ونشره في المجال العسكري.

3 - التفاهمات والمبادئ

بالاستناد إلى الإطار القانوني المبين أعلاه وانطلاقاً منه، ومع مراعاة الشواغل الإنسانية والأخلاقية والأمنية والشواغل المتعلقة بالاستقرار الاستراتيجي أيضا، ينبغي مواصلة تطوير التفاهمات والمبادئ التالية:

1 - مسؤولية البشر وخضوعهم للمساءلة ومشاركتهم

- **المسؤولية والمساءلة** – يجب على الدول أن تضمن أن يظل البشر مسؤولين وخاضعين للمساءلة في جميع الأوقات، وفقاً للقانون الدولي الساري، عن القرارات التي تتضمن استخدام الذكاء الاصطناعي في المجال العسكري.
- **التحكم والتقدير البشريين المناسبين للسياق** – إن القرارات العسكرية الحاسمة، بدءا من غرفة الاجتماعات وانتهاء بساحة المعركة – ولا سيما تلك التي تنطوي على استخدام القوة، يجب أن تُتخذ دائمًا باستخدام التحكم والتقدير البشريين المناسبين للسياق. ويمكن للذكاء الاصطناعي في المجال العسكري أن يساعد في اتخاذ القرارات، لكن ينبغي ألا يحل محل الاعتبارات والتقديرات القانونية والأخلاقية، كأن تكون له مثلا استقلالية ذاتية معرفية في اتخاذ القرارات. ويجب على الدول ألا تدمج هذه المنظومات إلا في سلسلة القيادة والتحكم التي يكون فيها البشر قادرين على الحفاظ على التقدير وعلى ممارسة مستويات مناسبة من التحكم. وينبغي معالجة التحيزات غير المقصودة قدر الإمكان.

2 - الموثوقية، وقابلية التنبؤ/قابلية التفسير، والمتانة

- **الموثوقية** – يجب أن يكون الذكاء الاصطناعي في المجال العسكري موثوقًا به لمنع العواقب غير المقصودة أو الأعطال، لا سيما إذا كان من المحتمل أن يكون له تأثير سلبي أو أن يضر بالمدنيين والأعيان المدنية. ويجب ألا يُستخدم الذكاء الاصطناعي في المجال العسكري إلا إذا كان من الممكن التنبؤ بآثاره وعواقبه بشكل معقول.
- **قابلية التنبؤ/قابلية التفسير** – يجب أن تكون عمليات اتخاذ القرار في الذكاء الاصطناعي قابلة للتنبؤ بها وقابلة للتفسير من قبل المسؤولين عن نشرها، مما يسمح لهم بفهم واستباق سلوكيات المنظومة.
- **المتانة** – يجب أيضا أن يكون الذكاء الاصطناعي في المجال العسكري متينا – على الصعيدين التقني والتشغيلي – حتى يظل آمنًا ومأمونًا عند نشره واستخدامه.

3 - الحد من المخاطر

- **تعزيز الإلمام بالحالة** – ينبغي استخدام الذكاء الاصطناعي لتحسين الإلمام بما يدور في ساحة المعركة من خلال جملة أمور منها الكشف عن وجود المدنيين بهدف الحد من احتمال وقوع الضرر.

- **التحليلات التنبؤية** – ينبغي استخدام النماذج التنبؤية القائمة على الذكاء الاصطناعي للمساعدة في تقييم المخاطر ووضع استراتيجيات الحد من تصعيد النزاعات، من بين أمور أخرى، ومنع وقوع إصابات في صفوف المدنيين.
- **ضوابط الحماية المدمجة** – ينبغي أن يشتمل الذكاء الاصطناعي في المجال العسكري على ضمانات تقلل من الضرر وتسمح بالتدخل البشري المناسب في حال حدوث أعطال في المنظومة.

4 – تجنب المسارات الجديدة للتصعيد

- **الاستقرار** – يجب ألا يُصمم الذكاء الاصطناعي في المجال العسكري أو يُنشر أو يُستخدم إلا بطريقة لا تؤدي إلى تأجيج التوترات الدولية أو إيجاد مسارات جديدة نحو التصعيد.
- **تحديد الأسلحة** – يمكن للذكاء الاصطناعي أن يدعم تحديد التسلح ويجب ألا يقوض القواعد والصكوك القائمة المتعلقة بعدم الانتشار وتحديد الأسلحة ونزع السلاح، كما يجب ألا يعرقل الامتثال لهذه القواعد، لا سيما فيما يتعلق بالأسلحة البيولوجية والنووية.
- **إدارة الأزمات** – يمكن للذكاء الاصطناعي في المجال العسكري أن يدعم تخفيف التصعيد وإدارة الأزمات.

5 – إدارة دورة حياة منظومات الذكاء الاصطناعي العسكرية

يتطلب الاستخدام العسكري المسؤول للذكاء الاصطناعي اتباع نهج شامل ومراع للمخاطر يتناول دورة حياة الذكاء الاصطناعي بأكملها في المجال العسكري. ويشمل ذلك تصميم هذه المنظومات وتطويرها واختبارها ونشرها وتشغيلها وتحديثها وإخراجها عن الخدمة. وفي كل مرحلة، يجب مراعاة الاعتبارات القانونية والإنسانية والتشغيلية والفنية ذات الصلة بشكل منهجي. ويشكل هذا النهج القائم على دورة الحياة أمرا ضروريا بشكل خاص حين يتعلق الأمر بالذكاء الاصطناعي العالي المخاطر في المجال العسكري، مثل الذكاء الاصطناعي الذي يُستخدم في الأسلحة الذاتية التشغيل أو في اختيار الأهداف أو في المساعدة في عملية اتخاذ قرار قد ينجم عنها إلحاق الضرر أو الموت بالأشخاص أو الإضرار بالأجسام، وبشكل عام في المجالات التي تخضع فيها القرارات للقانون الدولي الإنساني. وبالنسبة إلى المنظومات المنخفضة المخاطر، مثل أدوات الدعم الإداري أو نظم التخطيط اللوجستي، ينبغي تطبيق إدارة دورة الحياة على أساس تقييم المخاطر الملائم للسياق.

- خلال مرحلة التصميم والتطوير، يجب أن تضمن الدول تدريب المنظومات على مجموعات بيانات تمثيلية عالية الجودة – يوجد بها أقل قدر ممكن من التحيزات – مما يتيح استخدامها بالامتثال الكامل للقانون الدولي والقواعد والمعايير الدولية، لتقليل التحيز غير المرغوب فيه.
- في مرحلة الاختبار والتقييم، يجب تنفيذ إجراءات صارمة للاعتماد والتحقق من أجل التأكد من الموثوقية والامتثال القانوني والمتانة التشغيلية في ظل ظروف واقعية.
- في مرحلة النشر والاستخدام العملياتي، يجب أن تكون هناك ضمانات لرصد أداء النظام وضمان التحكم والتقدير البشريين المناسبين للسياق وتمكين التدخل البشري الملائم.

- خلال مرحلتي التحديث والتعلم، يجب أن تضع الدول بروتوكولات صارمة فيما يتعلق بتعديلات المنظومة، بما في ذلك التحكم في الإصدارات، وعمليات لإعادة الاعتماد والموافقة الرسمية.
- بالنسبة إلى مرحلة السحب أو الإخراج من الخدمة، يجب أن تكون هناك تدابير تمكن من تعطيل المنظومات أو أرشفتها بشكل آمن لمنع إساءة استخدامها أو تفعيلها أو إعادة نشرها عن غير قصد.

4 - الحوكمة الدولية

تؤكد سويسرا على أهمية أن تقود الأمم المتحدة عملية مستدامة وشاملة للجميع بهدف ترسيخ فهم مشترك لفوائد الذكاء الاصطناعي في المجال العسكري، وما ينطوي عليه من مخاطر وتحديات، وبهدف وضع مبادئ لاستخدامه بشكل مسؤول. وبناءً على ذلك، لا بد من إشراك جميع الدول الأعضاء وأصحاب المصلحة المعنيين، فضلاً عن ممثلي قطاعات العلوم والتكنولوجيا والمجتمعات المدنية والأوساط الأكاديمية، لضمان الشرعية والخبرة والدعم الواسع النطاق. وينبغي أن تكون عمليات الأمم المتحدة ذات الصلة شفافة، وأن تُعقد بانتظام، وأن تتسق مع المبادرات الأخرى ذات الصلة.

ويجب أن يكون الهدف الأسمى لجميع جهود الحوكمة الدولية للاستخدام المسؤول للذكاء الاصطناعي في المجال العسكري هو ضمان الامتثال للقانون الدولي، ولا سيما القانون الدولي الإنساني. وبالإضافة إلى ذلك، يجب أن تكون الشواغل الإنسانية والأخلاقية والحفاظ على الاستقرار والحد من المخاطر الأمنية في صميم هذه الجهود. وينبغي أن تساعد أطر الحوكمة الفعالة والمعايير المشتركة والحوار المستمر المتعدد الأطراف في منع التصعيد غير المقصود، وتعزيز الشفافية والثقة المتبادلة، وتقوية دور القانون الدولي في أوقات الاضطرابات الناجمة عن التكنولوجيا. وستساهم الدول في تهيئة بيئة أمنية أكثر قابلية للتنبؤ وأكثر مرونة وسلاما إن رسخت حوكمة الذكاء الاصطناعي في المجال العسكري في هذه المبادئ.

ويمكن أن تشمل الجهود المحددة ما يلي:

- تعزيز الفهم المشترك أو التعريفات والنطاق والمصطلحات المشتركة المتعلقة بالذكاء الاصطناعي في المجال العسكري
- تحديد الفرص والشواغل الإنسانية والقانونية والأمنية والأخلاقية وفهمها بشكل أفضل
- استكشاف تدابير الشفافية وبناء الثقة
- وضع المبادئ والمعايير وأفضل الممارسات وتوصيات أخرى
- تقديم توجيهات من أجل تنفيذها

أوكرانيا

[الأصل: بالإنكليزية]
[11 نيسان/أبريل 2025]

ظلت أوكرانيا تعمل بنشاط على تطوير الذكاء الاصطناعي وتطبيقه في مختلف مجالات الأنشطة، بما في ذلك المجال العسكري. وتدرك أوكرانيا بوضوح الإمكانات التي تنطوي عليها هذه التكنولوجيا لتعزيز رفاه الإنسان والقدرات العسكرية، والمخاطر الكبيرة الناتجة عن إساءة استخدامها في المجال المدني ولا سيما

في المجال العسكري. وتشتد هذه المخاطر بشكل خاص في سياق الغزو الشامل لأوكرانيا الذي قام به الاتحاد الروسي بدون سابق استفزاز وبلا مبرر، والذي انتهك خلاله بشكل منهجي قوانين وأعراف الحرب والقانون الدولي الإنساني.

وتدعم أوكرانيا الجهود الدولية الرامية إلى بناء توافق عالمي في الآراء بشأن التطوير والنشر والاستخدام المسؤول للذكاء الاصطناعي في المجالين المدني والعسكري، وتشارك في تلك الجهود.

وحتى تاريخه، وقعت أوكرانيا، في جملة أمور أخرى، على إعلان بليتشلي في عام 2023؛ وهي إحدى الدول المؤيدة للإعلان السياسي بشأن الاستخدام العسكري المسؤول للذكاء الاصطناعي وخاصية التشغيل الذاتي، الذي صدر في قمة تسخير الذكاء الاصطناعي المسؤول في المجال العسكري لعام 2023، المعقودة في لاهاي؛ وهي دعمت الدعوةَ إلى العمل من أجل الذكاء الاصطناعي المسؤول في المجال العسكري التي جرى الاتفاق عليها في قمة تسخير الذكاء الاصطناعي المسؤول في المجال العسكري لعام 2023 ومخططَ العمل من أجل الذكاء الاصطناعي المسؤول في المجال العسكري الذي اعتُمد ليكون وثيقة ختامية لقمة تسخير الذكاء الاصطناعي المسؤول في المجال العسكري لعام 2024؛ وانضمت إلى البيان بشأن تسخير الذكاء الاصطناعي المستدام والشامل للجميع من أجل البشر والكوكب في قمة العمل في مجال الذكاء الاصطناعي لعام 2025 في باريس؛ وشاركت في تقديم جميع قرارات الجمعية العامة الثلاثة بشأن الذكاء الاصطناعي التي اعتُمدت حتى تاريخه، بما في ذلك القرار 239/79 بشأن الذكاء الاصطناعي في المجال العسكري وآثاره على السلام والأمن الدوليين.

وتقف أوكرانيا على أهبة الاستعداد للمشاركة الفعالة في المبادرات العالمية الجديدة الرامية إلى تشجيع التطوير الآمن والأخلاقي والمسؤول للذكاء الاصطناعي. وهي تدعم أيضا إجراء مناقشات في منظومة الأمم المتحدة، بما في ذلك داخل مجلس الأمن، بشأن الذكاء الاصطناعي بمختلف جوانبه.

ولما كانت أوكرانيا دولة محبة للسلام ليس لديها مطالبات إقليمية ضد الآخرين، وضحية للعدوان العسكري الروسي، ولا تعترف بأي مطالبات من هذا القبيل ضدها، فإنها تطور وتستخدم الذكاء الاصطناعي العسكري حصرا لتعزيز قدراتها الدفاعية من خلال ممارسة حق الدفاع عن النفس الذي ينص عليه ميثاق الأمم المتحدة.

وفي سياق استخدام الذكاء الاصطناعي في السياق العسكري، حددت أوكرانيا المخاطر الرئيسية التالية على السلام والأمن الدوليين:

- قد يؤدي التنافس في دمج الذكاء الاصطناعي في منظومات القتال والأسلحة إلى إطلاق جولة جديدة أكثر خطورة من سباق التسلح العالمي على حساب تحقيق أهداف التنمية المستدامة، ولا سيما ظهور منظومات أسلحة ذاتية التشغيل بالكامل تعمل دون تدخل بشري.
- كما هو الحال مع التكنولوجيات الرقمية الأخرى، بالتزامن مع تنامي خطر الهجمات السيبرانية وتزايد درجة تعقيدها وتوسع مجالات تطبيقها، أصبح الذكاء الاصطناعي في المنظومات العسكرية أكثر عرضة للتدخل السيبراني وللتلاعب به من قبل طرف مغرض يسعى إلى حرمانه من خصائص تطبيقاته المقصودة ووظائف استخدامه الانتقائي.

- يمكن أن يؤدي الاعتماد المفرط على الذكاء الاصـــطناعي في اتخاذ القرارات إلى فقدان التحكم البشري في العمليات العسكرية الحاسمة.
- قد يؤدي الدمج المتسرع للذكاء الاصطناعي غير مكتمل التطوير في منظومات الأسلحة، لا سيما عند وجود خلل في قدرات تحديد الأهداف، إلى آثار عشـوائية وإلى زيادة الخسـائر في صـفوف المدنيين.
- لا يوجد حاليًا إطار متعدد الأطراف لضبط انتشار الأسلحة المدمجة بالذكاء الاصطناعي.
- استخدام الأسلحة المدمجة بالذكاء الاصطناعي دون الالتزام بقوانين وأعراف الحرب والقانون الدولي الإنساني يثير شواغل قانونية وأخلاقية خطيرة.

المملكة المتحدة لبريطانيا العظمى وأيرلندا الشمالية

[الأصل: بالإنكليزية]
[11 نيسان/أبريل 2025]

الذكاء الاصـــطناعي مجموعة تكنولوجيات عامة الغرض بإمكان أي منها تمكين الآلات من أداء المهام التي يتطلب أداؤها عادةً ذكاءً بشرياً أو بيولوجياً، لا سيما عندما تتعلم الآلات من البيانات كيفية القيام بتلك المهام. وتشـهد تكنولوجيات الذكاء الاصـطناعي حاليا نضـجا متسـارعا ويتم تبنيها بوتيرة غير عادية. وباعتبارها مجموعة من التكنولوجيات ذات نظم وأساليب وتطبيقات مختلفة، فإنها تمر بمسارات تطور مختلفة وتترتب عليها آثار متباينة. وما لا ريب فيه هو قدرتها الكامنة على إحداث تحول جذري في كافة جوانب المجتمع والاقتصاد والسياسة، بما في ذلك مجالا الدفاع والأمن.

وترحب المملكة المتحدة بالفرصـــة التي يتيحها قرار الجمعية العامة 79/239 للنظر في الآثار المترتبة على الذكاء الاصـطناعي في المجال العسكري غير تلك المتعلقة بمنظومات الأسلحة الفتاكة الذاتية التشـغيل، التي كانت موضـع مناقشـات مسـتفيضـة وقيّمة، بما في ذلك المناقشـات الجارية في إطار فريق الخبراء الحكوميين المنشأ بموجب اتفاقية حظر أو تقييد استعمال أسلحة تقليدية معينة يمكن اعتبارها مفرطة الضـرر أو عشـوائية الأثر. وإن إجراء تقييم دقيق للآثار الاسـتراتيجية الأوسـع نطاقاً للذكاء الاصـطناعي العســكري، بالجمع بين الآراء والأفكار والممارســات الجيدة التي نوقشــت في المحافل الدولية الرســمية وغير الرسمية بشأن هذا الموضوع، سيسمح بإجراء مناقشة شاملة بشأن كيفية تحقيق أقصى استفادة من الفرص التي يتيحها الذكاء الاصطناعي في المجال العسكري، مع التصدي بفعالية في الوقت نفسه للمخاطر المرتبطة به.

الفرص التي ينطوي عليها الذكاء الاصطناعي في المجال العسكري

من المحتمل أن يؤدي إدماج الذكاء الاصطناعي في المجال العسكري إلى إحداث تحول في قطاع الدفاع وديناميات الأمن العالمي وطابع الحرب. وســتســاعد التكنولوجيات المتقدمة القائمة على الذكاء الاصطناعي، التي يمكنها تصنيف كميات كبيرة من البيانات من مصادر مختلفة وتنقيحها بشكل أسرع وأكثر شمولاً، على زيادة الكفاءة وتحسين عملية اتخاذ القرار، وتسريع وتيرة التخطيط العملياتي وزيادة دقته. ويمكن للذكاء الاصطناعي عند استخدامه في منظومات الاستخبارات والمراقبة والاستطلاع أن يوفر صورة أكثر دقة

عن السياق العملياتي وأن يمكّن المخططين من تقليل الأثر على المدنيين – مما يؤدي إلى زيادة حماية المدنيين والبنية التحتية المدنية. وستؤدي المهام اللوجستية الذاتية التشغيل ومهام التعامل مع الذخائر غير المنفجرة إلى الحد من الحاجة إلى وجود أفراد عسكريين على الأرض. وبالتالي، يمكن للذكاء الاصطناعي في السياق العسكري أن يعزز الأمن الوطني والدولي ويحد من المخاطر المحدقة بحياة البشر ويقلل من الخسائر البشرية.

وقد حددت الأبحاثُ التي أجرتها وزارة الدفاع في المملكة المتحدة بشأن الذكاء الاصطناعي وحفظ السلام الطرق التي يمكن أن تستفيد بها عمليات السلام من القدرات والمنظومات المعززة بالذكاء الاصطناعي، ويشمل ذلك ما يلي:

- القدرة التحليلية التي ستؤدي إلى تحسين الإلمام بالحالة واتخاذ القرارات العملياتية وتخطيط السيناريوهات والقدرة على تحليل المشاعر.
- يمكن للمنظومات الذاتية التشغيل، مثل المركبات الجوية غير المأهولة، أن توفر تغطية معززة للمناطق الجغرافية الشاسعة أو المناطق العالية المخاطر (التي قد يكون من الخطر على أفراد حفظ السلام الاحتفاظ فيها بوجود دائم).
- يمكن للخدمات اللوجستية أن تؤدي إلى تحسين تقديم الرعاية الصحية وتقديم المعونة للسكان المحليين، ودعم تحقيق أهداف البعثة وبناء ثقة المجتمع المحلي.

ويمكن تطبيقُ هذه القدرات لتعزيز رصد اتفاقات تحديد الأسلحة واتفاقات السلام والتحقق منها، مما يسهل الكشفَ عن الانتهاكات أو تأكيدَ الامتثال في الوقت المناسب وبطريقة ذات مصداقية. ويمكن لأدوات الذكاء الاصطناعي أن تمكّن من تحسين الكشفِ عن مختلف أنواع العمليات العدائية التي تكون دون عتبة النزاع المسلح، وتحديدِ تلك العمليات وإسنادها إلى المسؤول عنها والتحقق منها، مما سيؤدي إلى الحد من فعالية هذه الأنشطة وربما ردعها عن الحدوث أصلا. ويمكن لهذه الأدوات أن تساعد أيضًا في القيام في الوقت الحقيقي برصد وتحديد خطاب الكراهية على شبكة الإنترنت أو الدعاية أو التغيرات في المشاعر العامة التي قد تؤدي إلى تصعيد التوترات أو تقويض أي محادثات سلام أو وقف إطلاق نار.

التحديات والمخاطر

قد يؤدي استخدام الذكاء الاصطناعي في السياق العسكري إلى تفاقم المخاطر القائمة وقد يطرح تهديدات إضافية تكون فوق عتبة النزاع المسلح أو دونها. وقد يؤدي الاندفاع نحو تبني قدرات الذكاء الاصطناعي لاكتساب ميزة استراتيجية إلى استخدام البلدان للذكاء الاصطناعي بطرق غير مقبولة من الناحية القانونية أو الأخلاقية أو من ناحية السلامة. والمخاطر الجديدة المتمثلة في التصعيدات أو الحوادث الناجمة عن الذكاء الاصطناعي بسبب أعطال في منظومات الذكاء الاصطناعي أو بسبب ضعفها أو هشاشتها أو عدم نضجها أو عدم أمانها ستتطلب وضع بروتوكولات وآليات جديدة لتخفيف التصعيد. وقد تسعى الجهات الفاعلة المعادية إلى مهاجمة منظومات الذكاء الاصطناعي الوطنية وتقويض الثقة في أدائها وسلامتها وموثوقيتها (على سبيل المثال عن طريق "تسميم" مصادر البيانات، وإفساد مكونات الأجهزة داخل سلاسل الإمداد والتشويش على الاتصالات والأوامر)، مما قد يعطل تلك المنظومات ويشوه عملية صنع القرار العسكري في أوقات الأزمات وفي البيئات العملياتية الأخرى.

وفي أوقات النزاع، من المرجح أن تؤدي هذه التكنولوجيات – ونسق العمليات التي تتيحها – إلى تقليص شديد للوقت الذي يستغرقه اتخاذ القرار، وإرهاق حدود الفهم البشري، وقد تتطلب استجاباتٍ تكون بسرعة الآلة. ويعني الطابع المبهم الشبيه بالصندوق الأسود الذي يتسم به العديد من قدرات الذكاء الاصطناعي أن البشر كثيرا ما يكونون عاجزين عن فهم كيفية أو سبب إنتاجها لمخرجات معينة. وقد تؤدي العمليات القائمة على الذكاء الاصطناعي إلى سلوك غير متوقع وغير شفاف، وقد تجعل من الصعب التوصل إلى استنتاجات وتقديرات دقيقة بشأن نوايا الخصم، أو قد تسيء تفسير تلك النوايا أو تؤدي إلى عواقب غير مقصودة. وقد يضع المشغلون ثقة مفرطة في المخرجات الخوارزمية دون فهم كامل للافتراضات وأوجه القصور والعيوب الأساسية لمنظومات الذكاء الاصطناعي. وما لم توجد ضمانات ومعايير وبروتوكولات مناسبة، يمكن للمنظومات القائمة على الذكاء الاصطناعي أن تؤدي إلى تفاقم مخاطر سوء الفهم وسوء التقدير والتصعيد غير المقصود.

ومن المرجح أن يؤدي توافر قدرات أو أدوات الذكاء الاصطناعي المتقدمة وغيرها من التكنولوجيات ذات الاستخدام المزدوج على نطاق واسع إلى زيادة مخاطر الانتشار وتطوير أسلحة جديدة من قبل الجهات الفاعلة من الدول ومن غير الدول. ويمكن أن يُستخدم الذكاء الاصطناعي أيضًا في تضخيم أو تعزيز محاولات التضليل الإعلامي الهادفة إلى توليد العداء تجاه الدول، مما قد يتسبب في نشوب النزاعات وتصعيد التوترات.

التزام المملكة المتحدة بالذكاء الاصطناعي الآمن والمسؤول في المجال العسكري

تدرك المملكة المتحدة أن الذكاء الاصطناعي يثير مخاوف عميقة بشأن الإنصاف والتحيز والموثوقية وطبيعة المسؤولية والمساءلة البشرية، لا سيما في السياق العسكري. ومع أن لدى الدول تاريخا طويلا في اعتناق التكنولوجيات الجديدة ومع أنها ستستمر في الاعتماد على النظم القانونية ونظم السلامة والنظم التنظيمية المعمول بها منذ فترة طويلة، فإنه يجب علينا إدراك التحديات الخاصة الناشئة عن طبيعة الذكاء الاصطناعي وإدراك أهمية أن نثبت بشكل إيجابي أننا نتحلى بالمسؤولية وأننا جديرون بالثقة.

وتبين المملكة المتحدة التزامها بالذكاء الاصطناعي الآمن والمسؤول في 'استراتيجيتها الدفاعية في مجال الذكاء الاصطناعي' وفي 'المبادئ الأخلاقية للذكاء الاصطناعي' المتصلة بها التي أعدتها. وتبين هذه المبادئ الأخلاقية للذكاء الاصطناعي، المنصوص عليها في سياسة المملكة المتحدة "الطموحة والآمنة والمسؤولة"، اعتبارات الإطار الأخلاقي المتمثلة في محورية الإنسان، والمسؤولية، والفهم، والحد من التحيز والأضرار، والموثوقية. ويوفر منشور الخدمة المشتركة المعنون "الذكاء الاصطناعي القابل للتعويل عليه في مجال الدفاع" الذي نُشر في تشرين الثاني/نوفمبر 2024، توجيهاً واضحاً للأفرقة داخل وزارة الدفاع وخارجها بشأن كيفية تطبيق هذه المبادئ الأخلاقية للذكاء الاصطناعي من أجل تقديم خدمات وقدرات قائمة على الذكاء الاصطناعي تكون قوية وموثوقة وفعالة.

وتسعى المملكة المتحدة، من خلال المبادئ الأخلاقية للذكاء الاصطناعي التي أعدتها، إلى تعزيز الثقة في تكنولوجيات الذكاء الاصطناعي وتطبيقاتها، وتحقيق كامل إمكانات التعاون بين الإنسان والآلة، مع التخفيف في الوقت نفسه من المخاطر المرتبطة باستخدام الذكاء الاصطناعي أو إساءة استخدامه أو عدم استخدامه ومع منع العواقب غير المقصودة. ويتيح هذا النهج للمملكة المتحدة تسخير الابتكار والإبداع

الموجودين في قطاعي الدفاع والصناعة بما يمكن من التبني الطموح للحلول القائمة على الذكاء الاصطناعي.

وتوضح حكومة المملكة المتحدة أن أي استخدام من قبل المملكة المتحدة للذكاء الاصطناعي لتعزيز العمليات الدفاعية أو المنظومات أو القدرات العسكرية يخضع للقانون الوطني والدولي. وتسعى وزارة دفاع المملكة المتحدة دوماً إلى التقيد بالتزاماتها القانونية عبر تنفيذ مجموعة كاملة من الأنشطة بدءاً من قانون العمل إلى قانون الخصوصية والمشتريات وقانون النزاعات المسلحة، المعروف أيضاً بالقانون الدولي الإنساني. وتطبق الوزارة ممارسات وعمليات قوية لضمان التزام أنشطتها وموظفيها بالقانون. ويجري تطبيق هذه الممارسات والعمليات – وسيستمر تطبيقها – على القدرات القائمة على الذكاء الاصطناعي. ويجب أن يمتثل نشر القدرات القائمة على الذكاء الاصطناعي في النزاعات المسلحة امتثالاً كاملاً للقانون الدولي الإنساني، باستيفائها للمبادئ الأساسية الأربعة المتمثلة في التمييز والضرورة والإنسانية والتناسب. وموقفنا واضح وهو أن استخدام أي منظومة أو سلاح لا يستوفي هذه المبادئ الأساسية سيشكل انتهاكاً للقانون الدولي.

وتشكل المسؤولية والمساءلة البشرية التي تتم ممارستها من خلال المشاركة البشرية الملائمة للسياق أمرا بالغ الأهمية. وهذه المشاركة البشرية الملائمة للسياق ضرورية للوفاء بسياساتنا ومبادئنا الأخلاقية والتزاماتنا بموجب القانون الدولي الإنساني. وتختلف طبيعة المشاركة البشرية تبعا لطبيعة القدرات والبيئة العملياتية وسياق الاستخدام. وستضمن المملكة المتحدة أن يظل التحكم السياسي في أسلحتها النووية في أيدي البشر في جميع الأوقات.

مساهمة المملكة المتحدة في المبادرات الدولية

إن الاستقرار العالمي يتطلب تطوير الذكاء الاصطناعي العسكري بشكل طموح لكن مسؤول. ولا يزال فهم المجتمع الدولي للمخاطر والضمانات والمعايير المتعلقة باستخدام الذكاء الاصطناعي في السياق العسكري في تطور. وبالنظر إلى أن المخاطر دولية بطبيعتها، فإنها تتطلب استجابة عالمية.

وقد ظلت المملكة المتحدة في طليعة الجهود الدولية الداعمة للتطوير والاستخدام الآمن والمسؤول للذكاء الاصطناعي. وتفخر المملكة المتحدة بأنها استضافت مؤتمر القمة الافتتاحي المعني بسلامة الذكاء الاصطناعي الذي أقر إعلان بلتشلي بشأن سلامة الذكاء الاصطناعي، كما تفخر بدورها في إصدار التقرير الدولي عن سلامة الذكاء الاصطناعي – وهو أول مصنف توليفي شامل في العالم للأدبيات الحالية التي تتناول مخاطر منظومات الذكاء الاصطناعي المتقدمة وقدراتها، وقد نُشر في شباط/فبراير 2025، ويرسي تفاهمات مهمة لإثراء النقاش الدولي، مثل تسخير الذكاء الاصطناعي من أجل السلام والأمن. إننا ندعم الجهود المبذولة في إطار التعاهد الرقمي العالمي لسد الفجوات الرقمية وتعزيز الحوكمة الدولية في مجال الذكاء الاصطناعي لصالح البشرية.

إن المملكة المتحدة تدعم بنشاط المبادرات الدولية الرامية إلى الدفع قدما بالعمل فيما يتعلق بالمجال العسكري. وقد دعمنا عمل منظمات من قبيل مؤسسة راند – فرع أوروبا، وجامعة كاليفورنيا، بيركلي، والهيئة العالمية لتسخير الذكاء الاصطناعي المسؤول في المجال العسكري، على الجمع بين خبراء متنوعين ومشهود لهم بالخبرة لاستكشاف هذه القضايا، وفهم أحدث الأفكار، ورسم سبل للمضي قدمًا لصانعي السياسات مع تقديم توصيات قابلة للتطبيق.

وتواصل المملكة المتحدة مشاركتها النشطة في الحوارات الدولية بشأن مسائل الدفاع والأمن المتصلة بالذكاء الاصطناعي وتواصل تبادل خبراتها في تطوير وتفعيل نُهج آمنة ومسؤولة إزاء اعتماد الذكاء الاصطناعي في المجال العسكري. وترحب المملكة المتحدة بالتقدم الذي جرى إحرازه من خلال مبادرات مثل قمة تسخير الذكاء الاصطناعي المسؤول في المجال العسكري، التي شاركت المملكة المتحدة في استضافتها في عام 2024، والإعلان السياسي بشأن الاستخدام العسكري المسؤول للذكاء الاصطناعي وخاصية التشغيل الذاتي الذي قادته الولايات المتحدة، والذي يهدف إلى زيادة فهم الفرص والمخاطر الاستراتيجية وسبل معالجتها من خلال اتخاذ التدابير المناسبة التي تدعم الاستخدام الآمن والمسؤول للذكاء الاصطناعي. وتشكل أخلاقيات الذكاء الاصطناعي وضمان جودة الذكاء الاصطناعي مجالين ديناميين يتطلبان المشاركة والتعاون والتكرار المستمر.

آفاق المستقبل

تتطلع المملكة المتحدة إلى البناء على التقدم المحرز حتى الآن في العمليات القائمة، بما في ذلك من خلال المناقشات التي تدور في الأمم المتحدة استناداً إلى تقرير الأمين العام وتركز على إجراءات ملموسة. ونظراً لطبيعة الذكاء الاصطناعي في السياق العسكري، سيكون من الضروري اتباع نهج شامل للجميع ومتعدد أصحاب المصلحة، يستنير بالخبرة التقنية والعسكرية والقانونية من الدول والقطاع الصناعي والأوساط الأكاديمية والمجتمع المدني.

وعلى الرغم من وفرة المعلومات المتوفرة لدينا، فإن فهمنا الجماعي للتطبيقات والآثار العسكرية لا يزال متدنيا، ولا تزال هناك فجوات معرفية كبيرة وسوء فهم لطبيعة الذكاء الاصطناعي وقدراته. ويتعين الاضطلاع بمزيد من العمل لبناء قدرات الدول، وتعزيز فهمنا الجماعي لآثار الذكاء الاصطناعي العسكري ومخاطره وتحدياته المحتملة على المستوى الاستراتيجي، ووضع مصطلحات متفق عليها عالميًا لكي يتسنى إجراء مناقشات بناءة. وينبغي أن تركز المناقشات على التدابير والممارسات الملموسة والفعالة والملائمة التي بمقدورها أن تساعد على التصدي للمخاطر، بما في ذلك أمور مثل الضمانات وقواعد السلوك، وقنوات الاتصال الجديدة وآليات كفالة الشفافية الرامية إلى الحد من مخاطر سوء التفسير، وتحديث العقائد العسكرية وتدابير بناء الثقة واتفاقات تحديد الأسلحة على نحو يواكب أثر الذكاء الاصطناعي العسكري.

باء – الاتحاد الأوروبي

[الأصل: بالإنكليزية]
[11 نيسان/أبريل 2025]

يرحب الاتحاد الأوروبي بهذه الفرصة التي أتيحت له لتقديم وجهات نظره بشأن التحديات والفرص التي يطرحها الذكاء الاصطناعي في المجال العسكري على السلام والأمن الدوليين، وفقاً للقرار 239/79، الذي اعتمدته الجمعية العامة في 24 كانون الأول/ديسمبر 2024.

أولاً وقبل كل شيء، يود الاتحاد الأوروبي أن يذكّر بموقفه الثابت المتمثل في أن استخدام الذكاء الاصطناعي في المجال العسكري يجب أن يكون متوافقاً مع القانون الدولي، ولا سيما ميثاق الأمم المتحدة والقانون الدولي الإنساني والقانون الدولي لحقوق الإنسان.

وبالمثل، يود الاتحاد الأوروبي أن يذكّر بموقف آخر يتمسك به منذ أمد بعيد، وهو أنه يجب الإبقاء دائمًا على التقدير والتحكّم البشريين في استخدام القوة. ويجب أيضًا أن يظل الإنسان مسؤولًا وخاضعًا للمساءلة عندما يتعلق الأمر بالذكاء الاصطناعي في المجال العسكري لضمان تطبيق هذه التكنولوجيا بطريقة مسؤولة.

ويدرك الاتحاد الأوروبي أن تطبيق الذكاء الاصطناعي على المنظومات العسكرية ينطوي على فرص وتحديات في آن واحد. ويتسارع تطور الذكاء الاصطناعي بوتيرة مذهلة لا يمكن معها التنبؤ في هذه المرحلة الزمنية بجميع مزاياه أو مخاطره.

وفي هذا الصدد، يرحب الاتحاد الأوروبي بتركيز الأمم المتحدة المستمر على هذه المسألة وكذلك بالمناقشات الجارية في المحافل الدولية المعنية. وفي هذا الصدد، يقدّر الاتحاد الأوروبي على وجه الخصوص استمرار عملية الذكاء الاصطناعي المسؤول في المجال العسكري، التي بدأت في هولندا في عام 2023 مع أول قمة من قمم تسخير الذكاء الاصطناعي المسؤول في المجال العسكري، تلتها قمة تسخير الذكاء الاصطناعي المسؤول في المجال العسكري التي استضافتها جمهورية كوريا في عام 2024. ويرحب الاتحاد الأوروبي بمواصلة العملية، من خلال قمة تسخير الذكاء الاصطناعي المسؤول في المجال العسكري، المقرر عقدها في عام 2025 في إسبانيا، ويعرب عن امتنانه لإسبانيا لتنظيمها القمة المقبلة.

ويشير الاتحاد الأوروبي إلى أن النداء للعمل من أجل الذكاء الاصطناعي المسؤول في المجال العسكري لعام 2023، وكذلك مخطط العمل من أجل الذكاء الاصطناعي المسؤول في المجال العسكري لعام 2024، قد حظيا بتأييد جميع الدول الأعضاء في الاتحاد الأوروبي. ويعتقد الاتحاد الأوروبي أن المفهوم الذي انبثق عن قمة تسخير الذكاء الاصطناعي المسؤول في المجال العسكري ويتمثل في تنظيم عمليات شاملة يشارك فيها أصحاب المصلحة المتعددون بشأن مسألة الاستخدام العسكري المسؤول للذكاء الاصطناعي نهجٌ واعد. وفي هذا الصدد، يدرك الاتحاد الأوروبي قيمة المساهمات الأخرى التي قُدمت مؤخرًا مثل القمة الدولية للذكاء الاصطناعي والمؤتمر المعني بالذكاء الاصطناعي العسكري المسؤول الذي استضافته فرنسا في 10 شباط/فبراير من هذا العام. وينوه الاتحاد الأوروبي أيضاً بالعمل الذي جرى الاضطلاع به في إطار الإعلان السياسي بشأن الاستخدام العسكري المسؤول للذكاء الاصطناعي وخاصية التشغيل الذاتي، باعتباره مساهمة قيمة في النقاش الدولي الأوسع نطاقاً بشأن آثار الذكاء الاصطناعي على السلام والأمن الدوليين.

ويعتقد الاتحاد الأوروبي أن الوثائقَ الختامية الصادرة عن قمم تسخير الذكاء الاصطناعي المسؤول في المجال العسكري والإعلانَ السياسي، الذي وقعت عليه جميع الدول الأعضاء في الاتحاد الأوروبي، يكملان بعضهما بعضاً ويكتسيان أهمية كبيرة لمواصلة تطوير التفكير والحوكمة والحلول العملية على الصعيد العالمي للاستخدام العسكري المسؤول للذكاء الاصطناعي.

ويسلم الاتحاد الأوروبي بأن هناك مزايا عسكرية لتطبيق الذكاء الاصطناعي في المجال العسكري. ويصدق ذلك، على وجه الخصوص، على سرعة العمليات العسكرية وحجمها ودقتها. ويمكن للذكاء الاصطناعي أن يوفر ميزة تكتيكية فيما يتعلق بالإدارة والمعالجة المسبقة لمجموعات البيانات الهائلة الناتجة عن منظومات المراقبة والأسلحة والطائرات المسيرة وصور السواتل، وهو أمر يمكن أن يمكّن المشغلين البشريين من اتخاذ قرارات أسرع وأفضل. ويمكن لتطبيقات الذكاء الاصطناعي أن تؤدي إلى الحد من

التكاليف من خلال تحسين الخدمات اللوجستية أو صيانة المعدات عن طريق إدارة الصيانة بطريقة تنبؤية. وبالمثل، يمكن للذكاء الاصطناعي أن يتيح إجراء عمليات عسكرية على مسافات أبعد، كما يمكن له أن يزيد من دقة العمليات العسكرية في البيئات التي يسود فيها عدم اليقين.

وفي الوقت نفسه، فإن مزايا السرعة والحجم التي تتيحها تطبيقات الذكاء الاصطناعي في المجال العسكري تطرح بدورها تحديات. فالذكاء الاصطناعي يعمل على تسريع حلقة الرصد والتوجيه واتخاذ القرار والتنفيذ. وقد تؤدي زيادة قدرات السرعة والحجم إلى ظهور تصورات خاطئة ناجمة عن تباينات بين النوايا العسكرية والتحليلات التي تنتجها المنظومات القائمة على الذكاء الاصطناعي. ومن ثم، يمكن للذكاء الاصطناعي أن يساهم عن غير قصد في التصعيد. وتمثل السرعة أيضًا تحديًا أمام هدف الاحتفاظ بالتقدير والتحكّم البشريين في استخدام القوة.

وإزاء هذه الخلفية، يشدد الاتحاد الأوروبي على أهمية التعاون الدولي الذي يسعى إلى دراسة أثر الذكاء الاصطناعي في المجال العسكري وأطر الحوكمة الممكنة.

المرفق الثاني

Replies received from international and regional organizations, the International Committee of the Red Cross, civil society, the scientific community and industry[1]

A. International and regional organizations

African Commission on Human and Peoples' Rights

[11 April 2025]

I. Introduction

The African Commission on Human and Peoples' Rights (the African Commission), as the premier treaty-based human and peoples' rights body of the African Union (AU), is entrusted with the mandate of promoting and protecting human and peoples' rights in Africa under the African Charter on Human and Peoples' Rights (African Charter). In the African Commission's study on Addressing Human Rights Issues in Conflict Situations, the African Commission's Focal Point who led the study observed that 'it is … in conflict and crisis situations that the most egregious violations and abuses of rights are perpetrated…With the changes in the nature of conflicts and the attendant heightened threat to human and peoples' rights, there is a greater need for the human rights system to pay increasing attention to and provide effective responses to the challenges that these new dynamics present to the protection and observance of rights.' In the current context, one of the major new dynamics that carries serious implications for peace and security and therefore human and peoples' rights relate to Artificial Intelligence (AI) and in particular its rapid development and use in the military domain.

During its 1214th meeting, the AU Peace and Security Council (PSC), in requesting the AU Commission to conduct a study to assess the adverse impact of AI on peace and security, underscored the necessity of ensuring African perspectives in shaping global AI governance frameworks. Against this background and having regard to its work on AI and other technologies and human and peoples' rights[2] and human rights in peace and security, the African Commission is pleased to share its views in response to the invitation of the Secretary-General for submission of inputs on AI in the military domain and its implications for international peace and security.[3]

---

1 In accordance with operative paragraph 8 of General Assembly resolution 79/239, the replies received from international and regional organizations, the International Committee of the Red Cross, civil society, the scientific community and industry are included in the original language received. The Secretary-General remains committed to multilingualism as a core value of the United Nations.

2 Resolution ACHPR/Res. 473 (EXT.OS/ XXXI) 2021 on human and peoples' rights and artificial intelligence (AI), robotics and other new and emerging technologies in Africa, available at https://achpr.au.int/en/adopted-resolutions/473-resolution-need-undertake-study-humanand-peoples-rights-and-art.

3 The Focal Point of the African Commission on its study on human and peoples' rights and AI, robotics and other technologies acknowledges with appreciation the contribution of Professor Thompson Chengeta, who is the consultant providing technical assistance in the development of the study, through the Centre for Human Rights, University of Pretoria.

II. AI in the military domain and peace and security

The development and use of AI technologies in the military domain particularly to automate military functions such as surveillance, targeting, and the deployment of lethal force have far reaching consequences for peace and security and hence for human and peoples' rights. The AU Continental AI Strategy, endorsed during the 44th Extraordinary Session of the Executive Council of the African Union, highlights AI governance and regulatory challenges, particularly in military applications, warning that AI could exacerbate conflicts through inaccurate predictions or deployment of autonomous weapon systems. Additionally, the framework raises concern about disinformation, misinformation, cybersecurity threats, and military risks.

From the perspective of the development and use of AI in the military domain, peace and security should not be seen just from the perspective only of what it means for stability of states and societies. Beyond its conception under the UN Charter and public international law associated with friendly relations of states, peace and security is also a fundamental right of all peoples. The African Charter thus stipulates that 'All peoples shall have the right to national and international peace and security. The principles of solidarity and friendly relations implicitly affirmed by the Charter of the United Nations and reaffirmed by that of the Organization of African Unity shall govern relations between States.'[4]

The framing of peace and security as a right of peoples compels states to assess and govern the development and deployment of AI technologies in the military domain through a human rights lens that prioritizes the prevention of harm, suffering, and injustice. Together international law conception of peace and security, it places an affirmative duty on states to ensure that AI systems do not contribute to conflict, perpetuate structural inequalities, or violate the rights and dignity of individuals and communities. By embedding peace and security within the framework of human rights, states are not only accountable for avoiding direct acts of aggression, but also for proactively creating and maintaining environments in which human flourishing, security, and justice are protected from the potentially disruptive or harmful impacts of emerging military technologies.

The implication of AI in the military domain to peace and security, farmed comprehensively, thus goes beyond how it shapes the obligation of states for non-aggression. It also covers how algorithm-driven systems may dehumanize individuals, introduce bias, and lead to unaccountable or disproportionate harm. It raises critical questions about the erosion of human oversight, the potential for unlawful killings or violations of international humanitarian law, and the targeting of vulnerable or marginalized populations.

By transforming military capabilities, the application of AI in the military domain can also have implications for peace and security by heightening tendencies for engaging in hostilities. The resultant escalation of tension and violence will be inimical not only to stability and peace between and within states but also most importantly carries more adverse consequences for the development needs of the less developed parts of the world such as Africa. While AI may contribute to advancing the development needs of Africa, its development and use in the military domain can have devastating consequences for development detrimental in particular to the right to development enshrined in Article 22 of the African Charter.[5]

This link between peace and development is also central to the Sustainable Development Goals (SDGs), especially SDG 16, which promotes peace, justice, and strong institutions. Without peace and security, sustainable development cannot be achieved.

4 Article 23(1) of the African Charter.

5 All peoples shall have the right to their economic, social and cultural development with due regard to their freedom and identity and in the equal enjoyment of the common heritage of mankind.

Recognising this link is critical in the governance of military AI, as the militarisation of AI can aggravate instability, particularly in fragile regions, and undermine Africa's developmental aspirations. By reaffirming the interconnectedness of peace and development, the African Commission calls for a governance approach that upholds peace as both a human right and a developmental imperative.

III. The need for a human and peoples' rights-based regulation of the development and use of AI in the military domain

Given the ways in which the use of AI in the military domain transforms the conduct of hostilities and how the development of AI relies on the extraction of natural resources particularly critical minerals such as rare earth minerals, it is the submission of the African Commission that both the process of extraction of resources in the development of AI in the military domain and the use of AI in the military domain need to be in full compliance with human and peoples' rights standards and international law principles, including international humanitarian law.

First and foremost, it is of paramount significance that the development and use of AI in the military domain complies with the right to peace and security enshrined in Article 23 of the African Charter on Human and Peoples' Rights. As a right that is born out of the recognition of the inseparability of the enjoyment of other human rights states from peace and security, this right entails that the use of AI in the military domain should be consistent with the international law prohibition of the use of force enshrined in the UN Charter and the Constitutive Act of the African Union.

Second, the use of AI technologies in conflict settings need to ensure respect for applicable human and peoples' rights and international humanitarian law principles, including most notably needs to adhere to the principles of precaution, necessity, distinction, proportionality and legitimacy. These requirements apply irrespective of whether the context in which the use of AI in the military domain relates to international armed conflicts or non-international armed conflicts. As established in the African Commission's study,[6] parties to conflict are obliged to observe human rights standards where such conflicts do not meet the IHL threshold of armed conflict. As such, those who use AI technologies in conflict situations that do not meet the IHL threshold of armed conflict are legally obliged to respect and ensure respect for the human and peoples' rights standards established under treaty and customary international human rights law.

Third, the development of AI in the military domain and the use AI technologies in hostilities need to comply with the principle of transparency. This is fundamental because it is the basis for ensuring effective regulation of the development and use of AI in the military domain and for compliance with applicable human rights and international law standards. Additionally, transparency is critical for ensuring compliance with the obligation for respecting the dignity, privacy and data protection of individuals. The principle of transparency is also a pre-requisite for addressing some of the concerns that arise from use of AI in the military domain including bias (owing to the source and type of data used) and explainability. Transparency is also critical not only with the development of AI in the military domain but also with respect to the transfer of AI technologies in the military domain.

Fourth, from the perspective of human and peoples' rights and IHL, the other standard key to human rights and international law-based regulation of the development and use of AI concerns accountability. In the event of the occurrence of violations of human and peoples' rights standards or IHL principles from the development and use of AI in the

6 ACHPR, Addressing human rights issues in conflict situations, https://achpr.au.int/en/node/895.

military domain, there has to be both institutional and individual accountability. Accountability in this instance encompasses not only the measures that are taken against perpetrators but also the remedial steps that need to be put in place for redressing victims.

Firth, building and sharing of technical knowhow critical to ensuring regulation by states is the other principle. Recent developments including the jamming of GPS systems affecting flights reported in Eastern DRC and the deployment by the Islamic State of West Africa of armed drones, highlight not only the need for effective regulation but also the need for developing the requisite infrastructure and technical capacity for ensuring effective regulation.

IV. The link between the development of AI in the military domain and Africa's natural resources and its implications for peace and security

The African Commission is also of the view that when discussing peace and security, stakeholders must be aware of the link between development of military AI, Africa's natural resources – particularly critical minerals – and the notion of peace and security. Article 21(1) of the African Charter on Human and Peoples' Rights affirms: "All peoples shall freely dispose of their wealth and natural resources. This right shall be exercised in the exclusive interest of the people. In no case shall a people be deprived of it."[7] Article 21(5) further provides that "States parties to the present Charter shall undertake to eliminate all forms of foreign economic exploitation particularly that practised by international monopolies so as to enable their peoples to fully benefit from the advantages derived from their national resources."[8]

This provision is particularly important in the context of military AI, which depends heavily on critical minerals such as cobalt, lithium, and rare earth elements – resources abundantly found in Africa. The 2924 Report of the Chairperson of the African Commission's Working Group on Extractive Industries, Environment and Human Rights Violations, stressed the "significance of critical minerals for new and emerging technologies" and highlighted that Africa has been burdened by a "resource curse phenomenon."[9] The report of the Chairperson noted that "extraction of minerals and other resources not only fuels but also at times becomes the site where contestation over whose control and use triggers conflicts. In some instances, this has created a vicious cycle of insecurity and violence, a condition that not only leads to major human and peoples' rights violations but also the perpetuation of a vacuum of effective governance and the concomitant exploitative, socially and environmentally costly extraction of the resources of the continent."[10]

Therefore, governance of military AI must not only ensure the legal use of force but also address the exploitative chains of extraction that power such technologies. This requires strict oversight, equitable benefit sharing, and regional solidarity to prevent Africa's resources from being used to fuel further conflict and inequality.

7 Article 21(1) of the African Charter.

8 Article 21(5) of the African Charter.

9 African Commission's Working Group on Extractive Industries, Environment and Human Rights Violations (2024), https://achpr.au.int/en/intersession-activity-reports/extractive-industries-environment-and-human-rights-violations (accessed 08 April 2025).

10 As above.

V. Conclusion

The African Commission is of the view that the development and use of AI in the military domain carries far reaching consequences for international peace and security in general and for less developed parts of the world such as in Africa that historically suffered violations and remain vulnerable to the adverse impacts of the development and use of AI in the military domain without robust and effective legal regime for such development and use in the military domain. The African Commission affirms that the development and use of AI in the military domain needs to be regulated on the basis of international law, human and peoples' rights and international humanitarian law standards with particular regard to the development and peace and security interests and human and peoples' rights needs of less developed parts of the world.

More specifically, beyond and above the right to peace and security, the governance of AI in the military domain needs to ensure respect for applicable human and peoples' rights and international humanitarian law principles, including most notably needs to adhere to the principles of precaution, necessity, distinction, proportionality and legitimacy, the principles of transparency, accountability and redress for victims and the obligation to build and share technical knowhow necessary for enabling societies to avert the risks that the development and use of AI in the military domain carries for peace and security. Only by ensuring that the development and use of military AI are aligned with international legal standards including those relating to the right to peace and security, the right to development, the right to privacy and protection of personal data, the right to remedy and the responsibility for exercising human control, the right to and control over natural resources and by addressing the structural inequities underpinning global technological advancement, can states uphold their duties to their peoples and advance genuine peace, justice, and security in relation to the development and use of AI in the military domain.

B. International Committee of the Red Cross

[19 March 2024]

Summary

The full submission is available at: https://www.icrc.org/en/article/artificial-intelligence-military-domain-icrc-submits-recommendations-un-secretary-general.

The International Committee of the Red Cross (ICRC) welcomes the opportunity to submit its views for consideration by the United Nations Secretary-General, in accordance with resolution 79/239.

The recommendations that the ICRC makes in this submission are in line with its long-standing mandate and practice of promoting respect for and the development of IHL, including its application to new technologies of warfare. This submission is intended to support States in ensuring that military applications of AI comply with existing legal frameworks and, where necessary, identifying areas where additional legal, policy, or operational measures may be required.

1. Normative proposals: Reaffirming existing IHL as the starting point

The ICRC has consistently emphasized that, while IHL does not explicitly prohibit or regulate the use of AI in military applications, it does restrict its development and use, and

places strict constraints on AI when it is integrated into weapon systems or used in some way to conduct warfare.[1]

Existing and emerging normative proposals on the military application of AI should build upon established international legal frameworks and mechanisms, including IHL. Where necessary, these frameworks can be reinforced through the development of additional legal instruments, operational guidance or policy measures to address specific risks or challenges posed by emerging technologies. The form and content of such measures may vary depending on the specific use case. The ICRC encourages the international community to engage in concrete discussions on particular applications of AI in the military domain and to prioritize consideration of those that pose the greatest risks to people affected by armed conflicts.

2. A Human-centred Approach to military AI

In line with the resolution, the ICRC advocates for a human-centred approach to the development and use of AI in armed conflict.[2] This approach has at least two key dimensions: first, ensuring a focus on the humans who may be affected by the use of AI; and second, emphasizing the obligations and responsibilities of the humans using or ordering the use of AI in military operations.

Despite the growing development of AI-related technologies in the military domain, IHL requires individuals to make legal determinations. Humans must, for instance, determine the lawfulness of attacks that they plan, decide upon or execute, and they remain accountable for those determinations. The ICRC considers that human judgement is crucial for reducing humanitarian risks, addressing ethical concerns and ensuring compliance with IHL. Accordingly, while certain technical tasks may be carried out by machine processes, it is not the system itself that must comply with the law, but the humans using it.[3]

This does not mean that commanders and combatants cannot or should not use tools, including AI-decision-support systems. However, these tools must only be designed and used to support, rather than hinder or replace, human decision-making.[4] Further, States and parties to armed conflicts must ensure that human control and judgement are preserved in decisions that pose risks to the life and dignity of people affected by armed conflict. This is essential for ensuring respect for applicable laws, including IHL, and upholding ethical standards.[5]

3. Specific Applications of ai in the military domain

The ICRC has identified three specific applications of AI in the military domain that pose particularly significant risks to those affected by armed conflict:

1 This has also been affirmed by States, including in the UN General Assembly with Resolution 79/239.

2 ICRC, AI and machine learning in armed conflict: A human-centred approach, 2019 (updated in 2021).

3 ICRC, International Humanitarian Law and the Challenges of Contemporary Armed Conflicts: Building a Culture of Compliance for IHL to Protect Humanity in Today's and Future Conflicts (IHL Challenges Report), 2024, p. 61.

4 *Ibid*.; ICRC, IHL Challenges Report – Chapter 2: Contemporary and future challenges in the conduct of hostilities, 2019, p. 32.

5 ICRC, Decisions, Decisions, Decisions: computation and Artificial Intelligence in military decision-making, ICRC, 2024, p. 8.

1. *AI in Autonomous Weapon Systems*

Resolution 79/239 acknowledges the increasing integration of AI into weapons and weapon systems, a development that raises significant legal and humanitarian concerns. The integration of AI, particularly machine learning (ML) techniques, into autonomous weapon systems (AWS) exacerbates existing challenges posed by AWS in ensuring compliance with IHL. In particular, it increases difficulties for human users to understand, predict, and control the system's functioning and effects.

Users of AWS must be able to, with a reasonable degree of certainty, predict the effects of that weapon in order to determine whether it can be directed at a specific military objective, and take steps to limit those predicted effects, as required by IHL. This entails the ability to understand the functioning of the AWS: the nature and functioning of its sensors, the definition of its target profile and the potential effects in the circumstances of use, including any risk of error or malfunction. This is particularly relevant for AWS that function in opaque ways (the "black box" challenge), such as AWS relying on AI techniques, which prevent the human user from being able to understand, predict or explain the system's output. This impossibility effectively results in a lack of control over the weapon's effects, rendering it indiscriminate by nature.

In this regard, we reiterate the joint call made by the ICRC President, with the UN secretary-general,[6] for new, legally binding rules prohibiting certain AWS and constraining the use of others.[7] In particular, we recommend a prohibition on

- unpredictable autonomous weapons – those that, due to their design or the circumstances and manner of use, do not allow a human user to understand, explain or predict the system's functioning and effects;
- autonomous weapons designed or used to target humans directly. This is required because of the significant risk of IHL violations and the unacceptability of anti-personnel autonomous weapons from an ethical perspective.[8]

The ICRC supports all efforts by States to urgently adopt a legally binding instrument to regulate AWS, in whichever forum they choose.[9] The integration of AI into AWS should also be considered when discussing normative proposals on military applications of AI. Doing so is essential to ensure a consistent and comprehensive approach to the regulation of military AI, to avoid normative gaps, and to effectively address the serious legal, ethical, and humanitarian risks that are exacerbated by the integration of AI into AWS. In this regard, the ICRC considers it important that binding prohibitions and restrictions on AWS, including AWS that incorporate AI, are integrated into broader discussions on the governance of military AI.

2. *AI in Military Decision-Making*

AI decision-support systems (AI-DSS) are computerised tools that bring together data sources – such as satellite imagery, sensor data, social media feeds or mobile phone signals – and draw on them to present analyses, recommendations and predictions to decision makers.

---

6 ICRC, Joint call by the United Nations Secretary-General and the President of the International Committee of the Red Cross for States to establish new prohibitions and restrictions on Autonomous Weapon Systems, 2023.

7 ICRC, ICRC Submission on AWS to the UN Secretary-General, 2024, p. 6.

8 *Ibid*.

9 *Ibid*.

The use of AI-DSS raise concerns related to system functioning, data quality, and human-machine interaction. These systems risk increasing the rate of unforeseen errors, perpetuating problematic biases – particularly those based on age, gender, ethnicity, or disability, and making it difficult for the users to understand how and why the system generates its output from a given input.

Generally, AI-based systems will perform better when given well-defined goals and access to representative and high-quality data. However, armed conflict environments are marked by uncertainty, volatility, and deliberate deception techniques by adversaries, which makes it extremely difficult to obtain reliable or transferable data. Even where good data exists, it may not reflect the specific operational or humanitarian dynamics of a particular context.[10] Moreover, for AI systems that rely on training data, the utility of those data can rapidly diminish once a conflict begins. Parties to armed conflicts will continuously seek to maintain the initiative and operate in a manner that is not anticipated by their adversary, adapting their strategies and tactics accordingly. This can fundamentally alter the environment in which the system was expected to operate, making the original data no longer representative of the new operational conditions. In such cases, the system's outputs may become unreliable, and the AI model may require re-evaluation or retraining in order to remain fit for purpose.

Human interaction with these systems raises further concerns, such as "automation bias" – a propensity to rely on machine outputs even when other available information may call those outputs into question – which is particularly pronounced in high-pressure or stressful environments like in armed conflicts.[11] Taken together, these factors can hamper a user's ability to scrutinize the information available. The practical consequence might be, for instance, that someone plans, decides upon or launches an attack based solely on an AI-DSS's output, thereby effectively serving as a human rubber stamp rather than assessing the lawfulness of the attack by considering all the information reasonably available including the AI-DSS output.[12]

On the positive side, the careful use of AI-based systems may facilitate quicker and more comprehensive information analysis, which can support decisions in a way that enhances IHL compliance and minimizes risks for civilians. In the context of urban warfare in particular, the ICRC has recommended that online open-source repositories should be used to gather information about the presence of civilians and civilian objects.[13] Importantly, IHL imposes obligations to take constant care to spare the civilian population and to take all feasible precautions in attack. Therefore, in developing and using AI-DSS, armed forces should be considering not only how such tools can assist them to achieve military objectives with less civilian harm, but also how they might be designed and used specifically to protect civilians. However, the important point is that these computer outputs can inform but must not displace the need for legal determinations.

Beyond targeting decisions, militaries are also exploring the use of AI to support other operations traditionally carried out by humans, including detention operations. While technology deployed responsibly and with robust human oversight can contribute to IHL

10 ICRC, IHL Challenges Report, 2024, pp. 64–65; ICRC, AI and machine learning in armed conflict: A human-centred approach, 2019 (updated in 2021); ICRC, Decisions, Decisions, Decisions: Computation and Artificial Intelligence in Military Decision-Making, ICRC, 2024, pp. 31 and 54.

11 ICRC and the Geneva Academy, Artificial Intelligence and Related Technologies in Military Decision-Making on the Use of Force in Armed Conflicts: Current Developments and Potential Implications, ICRC, 2024, p. 17.

12 ICRC, IHL Challenges Report, 2024, p. 65.

13 *Ibid*., p. 66; ICRC, Reducing Civilian Harm in Urban Warfare: A Handbook for Armed Groups, 2023, p. 15.

compliance, it also carries risks including bias, lack of transparency, and faulty programming and analysis, all of which can undermine compliance with IHL.[14]

To support efforts by States and other actors to ensure that military uses of AI-DSS remain consistent with IHL and humanitarian principles, the ICRC has formulated a non-exhaustive set of preliminary recommendations relating to the development and use of AI-DSS in armed conflict. They focus on 1) ensuring human control and judgement; 2) system design requirements; 3) testing, evaluation, verification and validation; 4) legal reviews; 5) operational constraints on use; 6) user training; 7) after-action reviews; and 8) accountability, among others. The recommendations are annexed to the full version of this submission.

3. AI in Information and Communications Technologies

AI is expected to change how actors defend against and conduct information and communications technology (ICT) activities, including in armed conflict. In particular, States have noted with concern that the use of AI and other emerging technologies in malicious ICT activities may further increase their scale and speed, as well as the harm they may cause.[15] For example, AI enables tools to identify and develop exploits for new vulnerabilities in software or networks, or to conduct harmful ICT activities autonomously, whether in offence or in defence. The ICRC is concerned that this could increase the risks of indiscriminate attacks, incidental civilian harm, including damage to critical civilian infrastructure, as well as the uncontrolled escalation of conflict, particularly in complex and interconnected digital environments.[16]

Similarly, information or psychological operations are not a new feature of armed conflicts; however, AI is changing how information is created and spread. AI-enabled systems, particularly generative AI, have been widely used to produce harmful content – text, audio, photos and video – which is increasingly difficult to distinguish from authentic, original content.[17] The ICRC is concerned about the consequences for civilians that might result from the creation and spread of such information through ICT, including information that contributes to or encourages violence, causes lasting psychological harm, undermines access to essential services or disrupts the operations of humanitarian organizations.

In light of these concerns, the ICRC underlines the importance of applying existing international law, including IHL, to the use of AI in ICT activities. The ICRC urges States to ensure that the development and use of AI-supported ICT activities respect the protections afforded to civilians and civilian infrastructure in armed conflict. Moreover, in light of the emergence of increasingly autonomous ICT capabilities, the ICRC further encourages States to address the serious challenges posed by these tools, particularly by considering whether existing international law, including IHL, provides sufficient safeguards against the harm such tools can cause, or whether additional limits are needed.

4. Conclusion

The ICRC is grateful for the opportunity to share its above views and recommendations on ways to address the challenges and concerns raised by AI for the

14 ICRC, IHL Challenges Report, 2024, p. 22.

15 34th International Conference of the Red Cross and Red Crescent, Resolution 2 "Protecting civilians and other protected persons and objects against the potential human cost of ICT activities during armed conflict", 2024.

16 ICRC, IHL Challenges Report, 2024, pp. 66–67.

17 *Ibid*., pp. 58–59.

secretary-general's consideration, and stands ready to contribute further to assist States in taking effective action to address the risks posed by AI applications in the military domain.

C. Civil society

Autonorms

[10 April 2025]

The following is the AutoNorms project's submission pursuant to Resolution 79/239 on "Artificial intelligence in the military domain and its implications for international peace and security" adopted by the United Nations General Assembly on 24 December 2024. The resolution requests the UN Secretary-General to seek views, including those of Member States, civil society, the scientific community and industry, on "opportunities and challenges posed by the application of artificial intelligence in the military domain, **with specific focus on areas other than lethal autonomous weapons systems**". The AutoNorms team welcomes the opportunity for representatives of academia to submit their views on this important and timely topic.

The AutoNorms project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 852123). Led by Professor Ingvild Bode and hosted by the Center for War Studies at the University of Southern Denmark, the project examines how the integration of artificial intelligence (AI) technologies into weapon systems and military targeting shapes international norms governing the use of force.[1]

Introduction

Over the past 2-3 years, the international debate about applications of AI in the military domain has been characterized by two significant, near-simultaneous changes. First, there has been a **move away from its predominant focus on autonomous or AI technologies in weapon systems** towards considering AI technologies across a wider range of military decision-making tasks, especially in relation to targeting. To reflect his move, this submission focuses on the employment of **AI-based decision support systems (AI DSS),** or systems that are meant to be used as tools to directly or indirectly inform the complex process of use-of-force decision-making, for example, by analyzing large volumes of data, recognizing patterns within the data, predicting scenarios, or recommending potential courses of action to human decision makers.

Second, there has been a **growing emphasis on human-machine interaction** in the context of using AI in the military domain.[2] This emphasis results from the broad recognition that, even when humans are 'in' or 'on' the loop of targeting decision-making, they need to exercise a sufficient level of oversight, control, and agency over the targeting process. Human oversight is a governance principle featuring prominently across various international initiatives, including A/RES/79/239. However, **dynamics of human-machine**

1 The members of the AutoNorms team are Professor Ingvild Bode, Dr Hendrik Huelss, Dr Anna Nadibaidze, Dr Guangyu Qiao-Franco, and Dr Qiaochu Zhang. The AutoNorms project is based at the Center for War Studies, University of Southern Denmark, Odense, Denmark. For more information, please visit our website: www.autonorms.eu.

2 Ingvild Bode and Anna Nadibaidze, "Symposium on Military AI and the Law of Armed Conflict: Human-Machine Interaction in the Military Domain and the Responsible AI Framework," *Opinio Juris*, April 4, 2024, https://opiniojuris.org/2024/04/04/symposium-on-military-ai-and-the-law-of-armed-conflict-human-machine-interaction-in-the-military-domain-and-the-responsible-ai-framework/.

interaction as part of the use of AI DSS both introduce new issues and solidify existing sets of challenges that require governance attention. Our submission highlights these challenges and the need to ensure the exercise of human oversight and agency throughout the full targeting decision-making spectrum. It is structured in three parts, starting with explicating challenges of human-machine interaction, then commenting on the relative under-development of the international debate about AI DSS, and finally, sketching a way forward.

Challenges of human-machine interaction in the use of AI DSS

The use of AI DSS involves various dynamics of human-machine interaction because military personnel such as operators and intelligence analysts routinely and increasingly interact with a network of AI systems throughout the targeting process. These interactions involve multiple challenges **which have the potential to affect the exercise of human agency**, or humans' capacity to understand a system's functions and its effects in a relevant context; deliberate and decide upon suitable actions in a timely manner; and act in a way where responsibility is guaranteed.[3]

Dynamics of human-machine interaction result in **distributed agency between humans and AI systems, where they are not separated into two distinct entities but rather form part of a socio-technical system**.[4] As part of this system, both sides may influence each other in different ways, which then translate into various forms of distributed agency located along a spectrum. In some instances, dynamics of human-machine interaction will offer more opportunities for exercising human agency in targeting decisions. In other instances, however, the humans involved in use-of-force decision-making will be more constrained in their ability to exercise agency.

For example, **humans' ability to exercise agency might be limited by cognitive biases such as automation bias or anchoring bias**. Humans could over-trust AI DSS even when knowing that there might be malfunctions or unintended errors involved, risking an overreliance on algorithmic outputs without engaging in the critical deliberations and assessments that are needed to exercise human agency, especially in critical targeting decisions that might inflict death, destruction, and severe harm. Such biases are typically exacerbated by the increased speed of AI-assisted military decision-making, especially in contexts where there are high levels of pressure to act rapidly. They can also be exacerbated by AI DSS that are used for prescription or recommendations, because such systems restrict the options or courses of action available to human decision makers.

Moreover, given that AI DSS are likely to be employed not individually but rather as part of a network of systems, the increased complexity of interactions can result in situations where humans act upon some outputs suggested by AI DSS, but do not overall exercise a high quality of agency. Due to these and many other concerns related to interactions between humans and AI DSS, **there is a need to further investigate challenges of human-machine interaction that result in AI DSS not positively 'supporting' humans but rather undermining humans' ability to exercise agency**.[5]

---

3 Anna Nadibaidze, Ingvild Bode, and Qiaochu Zhang, *AI in Military Decision Support Systems: A Review of Developments and Debates* (Odense: Center for War Studies, 2024), https://www.autonorms.eu/ai-in-military-decision-support-systems-a-review-of-developments-and-debates/.

4 Ingvild Bode, *Human-Machine Interaction and Human Agency in the Military Domain*, Policy Brief No. 193 (Waterloo, ON: Centre for International Governance Innovation, 2025), https://www.cigionline.org/publications/human-machine-interaction-and-human-agency-in-the-military-domain/.

5 Anna Nadibaidze, "Do AI Decision Support Systems 'Support' Humans in Military Decision-Making on the Use of Force?" *Opinio Juris*, November 29, 2024, https://opiniojuris.org/2024/11/29/do-ai-decision-support-systems-support-humans-in-military-decision-making-on-the-use-of-force/.

The risks of not addressing challenges of distributed agency are substantial. First, situations where humans are restricted in their exercise of agency **raise questions about compliance with international humanitarian law**, which requires that humans be held accountable and legally responsible for violations of legal principles. Although humans remain officially in control of the selection and engagement of targets, there are concerns about the exact role played by humans in context of using AI DSS in practice.

Second, these concerns also extend to the risk of **negatively affecting moral agency and responsibility in warfare**. Challenges of human-machine interaction that result in distributed agency would allow humans to feel less morally responsible for decisions that could affect other people's lives. They also risk making the human role a nominal, 'box-checking' exercise which can *de facto* be compared with AI DSS playing an 'autonomous' role because the human role is substantially reduced.

Third, there are **security and operational risks related to distributed agency dynamics**, especially when they give too prominent roles to AI DSS and algorithmic outputs. AI systems often malfunction, are trained on biased sets of data which do not apply beyond the training context or specific contexts of use, as well as integrate assumptions that might not be strategically or operationally beneficial.

Various types of biases, issues of trust, uncertainties, targeting and military doctrines, political and societal contexts in which AI DSS are used – all these aspects **can lead to dynamics of distributed agency which limit the exercise of human agency and prioritize algorithmic outputs**. It is important to investigate these dynamics and ensure that distributed agency provides more opportunities than limitations to human decision makers in warfare.

Relative under-development of the international debate on AI DSS

Despite increasing reports about the use of AI DSS in recent and ongoing armed conflicts, and the significant challenges and risks they pose to the effective exercise of human agency, **the international debate on human-machine interaction in the use of AI DSS remains insufficiently developed**, particularly within intergovernmental UN settings. Current discussions on AI in the military domain, including those within the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (GGE on LAWS), have focused on the use of AI at the tail-end of the targeting process, specifically autonomy and AI in weapon systems. This narrow focus risks overlooking or failing to address critical normative, legal, ethical, security, and operational risks that can proliferate and compound throughout the entire targeting decision-making process.

An increasing, albeit still limited, number of stakeholders are raising this issue at international multistakeholder forums, such as the Summits on Responsible Artificial Intelligence in the Military Domain (REAIM). Some international non-governmental organisations and research institutes – such as the International Committee of the Red Cross (ICRC), the Stockholm International Peace Research Institute (SIPRI), the UN Institute for Disarmament Affairs (UNIDIR), and the Asser Institute – have initiated discussions on challenges posed by AI in the military domain, beyond the issue of autonomy in weapon systems. Despite this progress, there remains **a clear need to develop a more comprehensive and inclusive international multistakeholder debate to guide the responsible development and deployment of AI DSS in military contexts.**

Way forward

In closing, we sketch three ways intended to move the international debate about applications of AI in the military domain forward:

1. **Increase awareness for the implications of practices of designing, developing, and using AI DSS**. States and other stakeholders across industry, civil society, and academia engaged in the governance, development, and use of AI DSS for military targeting must consider the implications of their practices. These practices influence what counts as 'appropriate' ways of considering and employing AI DSS and thereby shape what becomes the accepted, requisite quality of human oversight and agency exercised over the whole process of use-of-force decision-making. To increase such awareness, the debate pursuant to A/RES/79/239 at the UNGA First Committee should centrally focus on the issue of AI DSS in the military domain.

2. **Consistently map both 'best' practices and 'problematic' practices associated with the design, development, and use of AI DSS**. To get a better sense of the direction that the design, development, and use of AI DSS take, states and other stakeholders need to closely map their own (and others') practices. While there have been some limited efforts to exchange potential best practices, we also need to be attentive to practices with potentially problematic effects. This should encompass practices exercised across the full life cycle of AI systems from development to use and post-use review. Mapping such practices would offer stakeholders a better overview of which practices may be beneficial, i.e., provide opportunities for the exercise of human agency, and which practices may be problematic, i.e., limit the exercise of human agency, and therefore assess the desirability of particular practices.

3. **Pursue the debate on AI DSS within a multistakeholder format**. States should work with diverse stakeholders – including academics across social sciences and technical disciplines, civil society representatives, and international organizations – to develop normative guidance and regulation, especially regarding the human role in military decision-making. Moreover, top-down processes towards governing AI DSS should be accompanied by a bottom-up, standard-setting process focused on establishing operational standards. Such an inclusive approach could strike a balance between national security and humanitarian concerns, while reinforcing the need to ensure that humans can exercise agency in use-of-force decisions.

Global Commission on Responsible Artificial Intelligence (AI) in the Military Domain

[11 April 2025]

1. Introduction

The Global Commission on Responsible Artificial Intelligence (AI) in the Military Domain (GC REAIM) welcomes the opportunity to contribute to the United Nations Secretary-General's report pursuant to resolution A/RES/79/239.

GC REAIM recognises that military applications of AI present both opportunities and challenges for global peace and security. Accordingly, the establishment of responsible and ethical governance – consistent with States' obligations under applicable international law – is essential. The global community must take proactive steps to ensure that military AI is developed and deployed in a manner that de-escalates rather than escalates conflicts; respects and enhances, rather than compromises, the sovereignty and territorial integrity of

states; promotes rather than threatens the security and safety of civilians; constrains and supports rather than erodes the existing rules-based international order.

In line with GC REAIM's resolute commitment to advancing international governance efforts, this note outlines some of the – non-exhaustive – views expressed by GC REAIM Commissioners and Experts on the implications of AI in the military domain to peace and security. The views presented are general in nature and will be further elaborated in the forthcoming GC REAIM report. While the Commission plans to present substantive and actionable recommendations for stakeholders in September 2025, this note does not yet include concrete proposals. As discussions among Commissioners and Experts are still ongoing, it instead highlights some of the key opportunities, challenges, benefits and risks posed by AI in the military domain to peace and security.

2. Technological Foundations

GC REAIM holds that meaningful policy deliberations on AI in the military domain must be grounded in a shared, foundational understanding of the underlying technologies and their potential trajectories. The complexity of AI technologies often gives rise to misunderstandings, inflated expectations, or misguided applications. Consequently, it is imperative to demystify AI through formal and well-defined frameworks that distinguish between current capabilities and speculative future developments. To support this objective, GC REAIM is developing a taxonomy which seeks to map the full spectrum of AI applications across military and broader peace and security contexts. The taxonomy differentiates between the implications of AI in operational activities – such as warfighting and intelligence – and administrative activities – such as logistics and personnel training and helps identify the specific applications of AI that should be prioritised in governance deliberations.

In its approach to the creation of a taxonomy, GC REAIM highlights the need for and contributes to a concerted effort to clarify, standardise, and encourage the accurate use of technical language with different layers of abstraction for policymakers, experts, and the public, thereby enhancing transparency, mutual understanding, and public trust. GC REAIM also cautions against the uncritical multiplication or adoption of new terminologies in AI governance discourse, unless these are clearly defined; and to ensure such terms are not used to circumvent or obscure existing legal obligations. Precision and consistency in language are the basis of responsible AI governance.

3. Implications for Peace, Security, and Stability

GC REAIM recognises that the integration of AI into the military domain presents benefits as well as both foreseeable and unforeseeable risks to international peace and security. A balanced approach to the range of opportunities and challenges emerging throughout the AI life cycle lies at the core of GC REAIM's method and is essential for responsible AI governance.

AI in the military domain may contribute to international peace and security in several important ways. At the developmental stage, the advancement of military AI capabilities may act as a deterrent to violence, as the mere development and presence of advanced technologies by responsible actors can encourage restraint by aggressors. Military AI may enhance early warning systems, strengthening conflict prevention strategies, and supporting arms control verification through AI-driven tools that foster transparency, trust, and cooperation among states – fundamental elements in conflict prevention. AI can also bolster national security and defence by improving the precision, accuracy, and efficiency of intelligence analysis and situational awareness, enabling real-time threat detection, and facilitating more efficient counterterrorism operations through predictive analytics and

autonomous systems. AI-powered systems can rapidly process vast amounts of complex data, enabling military forces to make timely, informed decisions that may prevent escalation and support conflict de-escalation efforts. These traits can also help improve targeting accuracy and precision, potentially reducing the risk of collateral damage or fratricide – attacks on one's own forces – and aiding compliance with International Humanitarian Law (IHL) to protect the security of protected persons, such as civilians and non-combatants, during armed conflict. Military AI may also reduce certain forms of human bias and enhance accountability by providing precise data, surveillance, and real-time monitoring, enabling clear attribution of actions to specific actors. In these ways, AI offers meaningful opportunities to reinforce adherence to international law and ethical standards, strengthening the normative foundation of the rules-based international order underpinning global peace and security.

AI in the military domain also presents a range of risks. As with the development of other general-purpose technologies, the development of AI in the military domain may accelerate arms races. AI technologies driven by the commercial market may be repurposed by militaries or soldiers in need or increase the access of violent non-state actors to AI-enabled military capabilities, which may intensify ongoing conflicts and contribute to broader instability. There are also concerns that states could employ AI technologies to suppress human rights, entrench internal repression, and destabilise both regional and global peace.

Concurrently, as with AI more broadly, the environmental consequences of military AI – such as the energy-intensive demands of AI systems, resource extraction, and ecological damage from AI-enabled military systems – could aggravate resource scarcity and environmental degradation, fuelling tensions and undermining long-term peace. However, given the impact militaries have on civilian technology development, efforts to reduce the environmental impact of AI in defence settings could have far-reaching beneficial consequences for all uses of AI. As such, considerations of environmental impacts should be a component of responsible AI governance in the military domain.

The large-scale data extraction required for AI development could intensify geopolitical rivalries, facilitate intrusive surveillance, and create distrust through opaque and exploitative data practices. Such deployment of military AI may perpetuate discrimination and exacerbate social divisions, undermining stability and ultimately international peace and security.

There are simultaneously significant concerns regarding the potential of integration of AI within the command, control, and communication (C3) structures of nuclear weapons. A number of Commissioners and Experts have emphasised that this is a red line that must not be crossed. The commitment of several nuclear-armed states to human decision-making surrounding the employment of nuclear weapons is therefore applauded. Further, the development of large-scale lethal autonomous weapon systems – such as swarms of anti-personnel devices – risks creating a new category of weapons of mass destruction, posing serious threats to global peace and security. Relatedly, AI may lower the barriers to creation and use of nuclear, chemical, or biological weapons by state or non-state actors, thus generating new challenges for arms control and non-proliferation regimes.

Beyond these strategic risks, AI may affect the character of war and lower the thresholds for armed conflict. By increasing the speed of armed escalation and driving changes in the capabilities of weapons systems, AI in the military domain may reduce states' confidence in their deterrent capabilities – particularly in the face of cyber infiltration risk – thus influencing how decision makers receive, process, and act on information. AI in the military domain could also exacerbate asymmetric warfare and violence by widening technological disparities that could increase the likelihood of force being used prematurely or disproportionately.

Operationally, inaccurate AI systems used for targeting can undermine the security of protected persons under IHL by increasing the risk of indiscriminate attacks, violations of proportionality, and failure to distinguish between combatants and civilians. Closely related to this is the risk of fratricide due to potential errors in target identification or decision-making, which can undermine operational effectiveness, escalate conflict, and erode trust within militaries and alliances. Finally, there are views that the use of certain AI systems in the military domain can create accountability gaps absent clear rules. By complicating the attribution of responsibility for unlawful actions, the deployment of AI in the military domain could undermine key principles of international law and state responsibility for internationally wrongful acts. This may complicate efforts to hold individuals or states responsible for violations, leading to a reduced deterrent effect against unlawful conduct. Without avenues to hold actors legally responsible, the enforcement of international law weakens, potentially destabilising peace, encouraging impunity, and exacerbating global insecurity.

4. Decision-Making and Responsibility

GC REAIM acknowledges the ethical and legal challenges that arise from integrating AI into military decision-making which may have a direct impact on preservation of peace and security. The relationship between human judgment and machine outputs is complex and without measures to ensure lawful, responsible and effective development and deployment, there can be an erosion of accountability and increased risks of unintended harm. As AI systems become more sophisticated and integrated within military capabilities, it is plausible that algorithmic decisions may become more commonplace across global battlefields, introducing moral and legal challenges regarding human control, oversight and judgment in diverse contexts.

To address these risks, GC REAIM promotes the need for context-appropriate human judgement over specific uses, capabilities and decisions of AI in military applications. The GC REAIM report will list considerations and conditions that underpin and support human responsibility, judgment and means of adequately evaluating relevant actions and decisions. This could include the introduction of technical standards for explainability, as well as maintaining appropriate human oversight in targeting decisions, assessments of precautions, proportionality and distinction, and other critical operational choices. However, given that the very definition of autonomy in machines suggests the minimisation or removal of the human, ensuring human responsibility and accountability may require focusing on human decision-making at earlier stages of a system’s life cycle, as the systems structure the behaviour of all who work with it. Human oversight is essential to uphold state obligations under applicable international law, in particular, IHL.

Military AI systems must be designed not only to support all individual and collective agents in the military domain to be effective in safely carrying out their lawful tasks, but also to do so responsibly and without compromising or undermining their status as moral human agents. GC REAIM suggests that military AI based socio-technical systems need to be explicitly and demonstrably designed to adequately attribute and apportion responsibilities and is determined to contribute to this process. For the security of protected persons, parties to armed conflicts should at all times be able to demonstrate that everything possible has been undertaken to create the conditions under which military personnel can effectively apply extant and widely shared principles and laws of armed conflict to their own situation, when using or relying upon AI components in the execution of their tasks.

5. Governance and Regulation

In light of both the opportunities and risks associated with military AI, GC REAIM supports a comprehensive governance framework that implements authentic international law. GC REAIM reiterates that existing legal regimes provide a solid foundation for regulating AI technologies. Governance must incorporate and account for procedural safeguards (due diligence and legal reviews, transparency of testing, evaluation, and validation, accreditation, and verification), substantive obligations drawn from various branches of international law, and soft law tools (military doctrines, national policies and strategies, norms and standards). In principle, all relevant international legal frameworks must be considered and applied. These include, but are not limited to, the following: (1) international law (*jus ad bellum*) which regulates when and how states use force, codifying a general prohibition on the use of force and exceptions such as in the case of self-defence, (2) international humanitarian law (*jus in bello*) which governs conduct during armed conflict and ensures the security of protected persons, (3) international human rights law.

GC REAIM further emphasises the critical role of international, regional, and domestic institutions in implementing and enforcing these legal norms. Effective governance requires collaboration across these levels and the inclusion of both binding (hard law) and non-binding (soft law) instruments. Soft law mechanisms, such as codes of conduct and ethical principles, can complement existing treaties and facilitate rapid, flexible responses to technological developments.

To address the diverse range of challenges surrounding the integration of AI into the military domain, GC REAIM supports proactive risk-mitigation and confidence-building measures. While binding regimes are challenging for general purpose technologies, there may be opportunities for rigorous monitoring, verification, and enforcement mechanisms inspired by successful global arms control regimes. For example, Commissioners and Experts have discussed ideas such as an Autonomous Incidents Agreement to reduce the risks of miscalculation among AI-enabled autonomous systems, or a committee or consortium that could set guidelines and recommendations surrounding the testing and evaluation of AI systems, including generative AI. GC REAIM also suggests that states and industries should consider adopting human-centred safety-by-design principles, implement red-teaming practices throughout AI system life cycles, and maintain clear chains of accountability for all actors. Only through robust multilateral dialogue and inclusive multi-stakeholder cooperation can AI be effectively governed to enhance peace and security rather than exacerbate global instability.

GC REAIM acknowledges that the development of a comprehensive governance framework for military AI faces several key challenges. First, there is the challenge of diverse interests and perspectives, with states, private companies, and civil society holding varying and sometimes conflicting views on the regulation of military AI. Second, the sensitivity surrounding national security and defence poses a significant barrier, as many states are reluctant to subject their military technologies to international scrutiny or regulation due to legitimate security interests. Third, achieving meaningful and substantive inclusivity in discussions is often difficult, as key stakeholders may be excluded or marginalised in decision-making processes. Fourth, a trust deficit between states, international organisations, and the private sector complicates efforts to establish cooperative governance. Fifth, the presence of crosstalk, incommensurability, and discursive dissonance arises due to the diverse backgrounds and expertise of stakeholders, making consensus-building challenging. Finally, these obstacles are compounded by the lack of clear frameworks that address the complex ethical, legal, and technical issues at the nexus of AI and the military domain. In light of these challenges, the final GC REAIM report will offer strategies to navigate and overcome these barriers in developing a robust governance framework.

6. Conclusion

GC REAIM observes that the rapid advancement and deployment of AI technologies in military contexts poses opportunities, challenges, benefits and risks for global peace and security. Balancing these considerations must be met with a technologically sound, inclusive, principled, and legally grounded approach to governance.

A clear understanding of AI's technological foundations is necessary to properly address its role in modern warfare. Ethical and legal responsibility should remain human-centred, and governance frameworks must rely on the robust application of international law, supplemented by cooperative multilateral efforts and soft law instruments when appropriate. In its formation and deliberations, GC REAIM has had the opportunity to reflect upon the conversations happening in broader governance processes, finding ways to effectively bridge gaps between disciplines and regional perspectives.

GC REAIM urges the United Nations and all State Parties to place these principles at the heart of global discussions on the implications of AI in the military and broader peace and security, for the present and future generations. Only through concerted international cooperation, guided by a shared commitment to human dignity, peace, and justice, can we ensure that the future of AI in the military domain is one that strengthens our common security.

InterAgency Institute

[11 April 2025]

The InterAgency Institute was established in December 2020 as a digital think tank, founded by expatriate and Global South women as a collective of researchers. It is in this condition that we address this submission on "opportunities and challenges posed to international peace and security by the application of artificial intelligence in the military domain, with specific focus on areas other than lethal autonomous weapons systems," following A/RES/79/239. With this, we seek to craft a complementary set of suggestions to develop the policy discussion in points where understanding that AI encompasses a wide array of data-processing techniques, and may be integrated into different types of warfare, in multiple parts of the organization, and at different levels.

The InterAgency Institute would like to point to overarching trends that fall within our areas of expertise, namely: (1) a focus on the global south, specially in how to prevent furthering the security gap; and (2) in how interagency cooperation in a time of greater mistrust may be leveraged to ensure the integration of AI in the military does not. Additionally, we make the point that Decision Support Systems (DSS) create analogous problems when compared to Autonomous Weapon Systems (AWS).

1. Addressing the security gap between the Global South and the Global North

The increasing technological intensity and digitalization of the battlefield are likely to increase the capacity gap between countries in the Global North & South. The "optimization of war" entails furthering this discrepancy, augmenting threats, and deteriorating the global security landscape. The wide range of AI-enabled solutions represents discrepant utility levels across tools.

While some tools require a low threshold (thus providing usually an equally low ceiling), the systems that pose the biggest military advantage require a high knowledge threshold to be implemented, therefore, will likely not be open source, and will only be available to entities with sufficient means to develop or acquire them. Given the experience

in past decades on multilateral forums, it is important to recognize that interest in access to these technologies will play a role in the negotiations.

In the long term, the current trend of "technological sovereignty" (or more specifically of restricted technological access due to global inequalities) may be transformed to undermine such technological control, creating far-reaching implications of this new revolution in warfare, involving stakeholders that may be reluctant to shape modern discussions due to a lack of current development of these technologies in their ecosystem.

2. InterAgency cooperation in times of distrust

These issues call for interagency cooperation at both the strategic and operational levels. The lack of interagency cooperation might lead to threat escalation and the eroding foundations for peace and security. Interagency cooperation should focus on formalizing specific channels for communication between different States, developing strategies for AI implementation that will not damage diplomatic relations, and generating more transparency in the interactions between agencies and contractors. The participation of different branches of government at the UN-level discussions is pivotal for a whole-of-government perspective in the deliberations. Beyond interagency cooperation at the governmental level, the wide array of applications of military AI calls for different sets of Confidence Building Measures (CBMs).

Since AI may be integrated in different warfare types and at different levels, its applications for different contexts have different ethical implications and consequences. Therefore, a monolithic understanding of risks posed by AI in the military context and consequently a unique set of CBMs would be inadvisable. CBMs for AI use in the strategic level of cyberspace will not be the same as CBMs for AI use in the tactical level of aerial warfare. Therefore, thinking about CBMs for military AI as a monolith will lead to inaccurate and in some cases inapplicable measures, undermining its effectiveness.

There is a necessity for sharing best practices in the introduction of AI into military procedures. In this sense, a trade-off should be made, prioritizing best practices that contribute to strengthening the aforementioned points of interagency cooperation and CBMs, and other practices that fall within the larger umbrella of strengthening international peace and security. Sharing of best practices relating to cybersecurity and reliability of the technology could also take place, but they should give priority to CBMs that focus on integration of AI at the strategic level and in manners that avoid the escalation of threats.

3. Decision Support Systems

Target identification or recognition via AI-enabled Decision Support Systems (DSS) entail analogue problems to Autonomous Weapon Systems (AWS). Digital dehumanization, lowering the threshold of violence, and automation bias are byproducts of that process that may only be avoided by the creation of red lines prohibiting such systems that replicate those concerns.

This problem stems not only from AI, but from a wider trend. Other data processing techniques that involve deterministic sorting of data that is not adequately processed by human operators also generate these problems. This caveat should be made to understand that not only systems with AI-enabled technology in DSS pose these kinds of threats, but a wider array of data gathering/processing techniques.

Conclusions and recommendations

- Formal interagency bodies to interface with multilateral AI/military tech negotiations
- Funding and support for academic research in the Global South focused on military AI implications;
- Regular technical-diplomatic summits focused on transparency, shared definitions, and threat perception;
- Prioritize capacity-building initiatives for Global South actors;
- Red lines and confidence building measures could be tailored to the specific technology and operational context;
- The discussions on Autonomous Weapon Systems encapsulate worries around AI-enabled Decision Support Systems. The creation of red-lines for these systems could benefit from building upon recommendations of the GGE on LAWS;

International Committee for Robot Arms Control

[11 April 2025]

The International Committee for Robot Arms Control (ICRAC) values the opportunity to submit our views to the United Nations Secretary-General in response to Resolution A/RES/79/239 "Artificial intelligence in the military domain and its implications for international peace and security."

Founded in 2009, ICRAC is a civil society organization of experts in artificial intelligence, robotics, philosophy, international relations, human security, arms control, and international law. We are deeply concerned about the pressing dangers posed by AI in the military domain. As members of the Stop Killer Robots Campaign, ICRAC fully endorses their submission to this report, and wishes to provide further detail regarding the concerns raised by AI-enabled targeting.

Increasing investments in AI-based systems for military applications, specifically AI-enabled targeting, present new threats to peace and security and underscore the urgent need for effective governance. ICRAC identifies the following concerns in the case of AI-enabled targeting:

1. AI-enabled targeting systems are only as valid as the data and models that inform them. 'Training' data for targeting requires the classification of persons and associated objects (buildings, vehicles) or 'patterns of life' (activities) based on digital traces coded according to vaguely specified categories of threat, e.g. 'operatives' or 'affiliates' of groups designated as combatants. Often the boundary of the target group is itself poorly defined. Although this casts into question the validity of input data and associated models, there is little accountability and no transparency regarding the bases for target nominations or for target identification. AI-enabled systems thus threaten to undermine the Principle of Distinction, even as they claim to provide greater accuracy.

2. Human Rights Watch research indicates that in the case of IDF operations in Gaza, AI-enabled targeting tools rely on ongoing and systematic Israeli surveillance of all Palestinian residents of Gaza, including with data collected prior to the current hostilities in a manner that is incompatible with international human rights law.

3. The increasing reliance on profiling required by AI-enabled targeting furthers a shift from the recognition of persons and objects identified as legitimate targets by their observable disposition as an imminent military threat, to the 'discovery' of threats

through mass surveillance, based on statistical speculation, suspicion and guilt by association.

4. The questionable reliability of prediction based on historical data when applied to dynamically unfolding situations in conflict raises further questions regarding the validity and legality of AI-enabled targeting.

5. The use of AI-enabled targeting to accelerate the scale and speed of target generation further undermines processes for validation of the output of targeting systems by humans, while greatly amplifying the potential for direct and collateral civil harm, as well as diminishing the possibilities for de-escalation of conflict through means other than military action.

Justification for the adoption of AI-enabled targeting is based on the premise that acceleration of target generation is necessary for 'decision-advantage', but the relation between speed of targeting and effectiveness in overall military success, or longer-term political outcomes, is questionable at best. The 'need' for speed that justifies AI-enabled targeting is based on a circular logic, which perpetuates what has become an arms race to accelerate the automation of warfighting. *Accelerating the speed and scale of target generation* effectively *renders human judgment impossible or, de facto, meaningless.* The risks to peace and security – especially to human life and dignity – are greatest for operations outside of conventional or clearly defined battlespaces. Insofar as the use of AI-enabled targeting is shown to be contrary to international law, the mandate must be to *not* use AI in targeting.

In this regard, ICRAC notes that the above systems present challenges to compliance with various branches of international law such as international humanitarian law (IHL), *jus ad bellum* (UN law on prohibition of use of force), international human rights law (IHRL) and international environmental law. In the context of military AI's implications for peace and security, *jus ad bellum*, a framework that prohibits aggressive military actions and regulates the conditions under which states may lawfully resort to the use of force, is the most relevant. In the same manner IHRL is important in this context because it is designed to uphold human dignity, equality, and justice – values that form the foundation of peaceful and secure societies.

International Humanitarian Law and Youth Initiative

[11 April 2025]

Artificial intelligence (AI) has gained a universal recognition during the 1950s'. Technological emergence has assisted humans in almost all facets of their lives thereby making work easier and faster. Moreso, the rapid growth of Artificial intelligence in technological field enthralling commercial investors, law makers, defense intellectuals and international competitors can be evidential in theoretical premises of international security. The use of Artificial intelligence (AI) in modern warfare particularly in the In the Middle East and North Africa, Ukraine/Russian armed conflict which has resulted in the killings of thousands of innocent civilians with women and children being the most vulnerable. The emergence of AI is expected to be utilized in improving all sectors in our daily lives However, its Negative application in the military domain continues to create Humanitarian crisis between warring parties making it of regional and international concern. The war in Gaza is one of the deadliest and most destructive war in history with technology playing a central role in enabling mass slaughter and destruction ranging from supplying the dystopian systems used to automate the killings and bombing.[1] Following the October 7 2023, there

1 Accessnow. (October 2024) Big Tech and the risk of genocide in Gaza: what are companies doing? Available at https://www.accessnow.org/gaza-genocide-big-tech/.

have been extensive reports evidencing the Israeli occupation forces use of surveillance technology, artificial intelligence, and other digital tool to determine who, what and when to attack in Gaza trip. Thus, this violates the principles of international humanitarian law which emphasize the necessity of distinguishing those in active combat and not[2] and to take necessary precautions when conducting an attack to minimize civilian harm.

IHLYI in this paper, responding to the request of the UN Secretary-General pursuant to a resolution A/RES/79/239, adopted by the General assembly on 24 December 2024 on Artificial intelligence in the military domain and its implication to international peace and security therefore, it analyzes AI In modern warfare, its implication to international peace and security and the role of technological companies in armed conflict.

Artificial Intelligence in Modern Warfare: A Legal and Humanitarian Perspective

The rules of international humanitarian law do not explicitly address the use of modern technological tools and artificial intelligence (AI) during armed conflicts. However, its core principles – such as distinction, proportionality, and precaution – remain applicable and binding on all parties. These principles require the differentiation between military objectives and civilians, and oblige parties to take all feasible measures to avoid or minimize harm to civilian populations. In recent years, militaries have contracted private companies to develop autonomous weapons systems. However, the armed conflict in Gaza stands out as one of the most prominent cases where commercially developed AI models – originally created in countries like the United States – have been employed in actual combat operations, despite the fact that these systems were not initially designed to make life-or-death decisions.

This shift highlights a troubling rise in the militarization of technology without clear legal or ethical oversight. While some of these tools may enhance operational efficiency, their unregulated use poses serious risks of human rights violations, especially amid a lack of transparency about how these tools function, the origin of the data they rely on, and the accuracy of their outcomes[3].

One of the most pressing concerns recently raised is the deployment of digital military tools based on unreliable data or flawed algorithms. Some of these systems depend on mass surveillance of Gaza's[4] population, including the collection of personal data prior to the outbreak of hostilities. Such practices raise legal and ethical questions regarding their compatibility with international obligations to safeguard privacy and prevent the misuse of personal information for the purpose of direct targeting.

Among the tools reportedly in use is a system that tracks population movement through mobile phone data to monitor evacuations from certain areas. Another generates lists of structural targets to be hit militarily. A third tool classifies individuals based on levels of suspicion regarding their affiliation with armed groups, while a fourth seeks to determine the precise location of a target in order to carry out a strike at the opportune moment. These tools largely rely on data extracted from mobile devices – whether through cell tower location information or GPS[5]. However, from a technical perspective, such data is

2 Article 48 of Additional protocol I of the Geneva convention.

3 Human Rights Watch, "Israel: AI-Powered Targeting Systems May Be Committing War Crimes in Gaza", 2024.

4 Associated Press, "Documents Reveal Israel's Use of AI Tools in Targeting Gaza", Investigative Report, 2024.

5 Human Rights Watch (2024). Questions and Answers: Israeli Military's Use of Digital Tools in Gaza Available at Questions and Answers: Israeli Military's Use of Digital Tools in Gaza | Human Rights Watch.

insufficiently precise to confirm an individual's presence at a specific location at a given time, particularly in conflict zones where individuals frequently change phones or numbers. Over-reliance on this technology may lead to fatal mistakes, especially when a mobile phone is used as a substitute for verifying a person's actual presence in a targeted area. Legally, the use of such systems without taking all feasible precautions to protect civilians constitutes a clear violation of international humanitarian law – particularly Article 57[6] of Additional Protocol I to the Geneva Conventions, which obliges parties to take constant care to spare civilian lives during military operations.

Given this reality, urgent questions must be raised about the future of AI in warfare and the legislative and legal mechanisms needed to regulate it. Without proper oversight, these tools risk becoming instruments of systematic human rights abuses rather than technologies aimed at ensuring greater protection for those affected by war.

Implications of Artificial Intelligence on International Peace and Security

Armed conflicts in various regions around the world, such as Gaza, Lebanon, Syria, Ukraine, and Libya, have had catastrophic humanitarian and security consequences. These conflicts have led to the mass displacement of civilian populations, depriving thousands of people of their basic rights such as food, water, shelter, and healthcare. These individuals live in dire humanitarian conditions, with a significant increase in deaths due to famine, thirst, and diseases caused by contaminated water, in addition to exposure to harsh weather conditions without protection.

In this context, the increasing use of artificial intelligence and drones as weapons in conflicts, particularly by Israel in the Gaza Strip[7], stands out. Since October 2023, there has been a notable escalation in the use of "quadcopters" to carry out precise and targeted strikes against civilians. These drones are equipped with data analysis algorithms and offensive capabilities, enabling them to target individuals based on tracking their movements or mobile phone signals.

According to documented reports, this technology has led to the death of more than 1,000 Palestinians by May 2024, including a significant number of women and children. This constitutes a grave violation of international humanitarian law, particularly Articles 51 and 57 of Additional Protocol I to the Geneva Conventions, which prohibit attacks on civilians and obligate parties to the conflict to take all necessary precautions to avoid harming them.

The concerns are not limited to the use of artificial intelligence against individuals but extend to the misuse of data. Relying on mobile phone tracking technologies (either through GPS data or cell tower signals) to pinpoint individuals' locations presents serious risks. Recent studies have shown that these systems do not provide enough accuracy to reliably determine someone's location, especially in conflict zones where phones may be swapped or disconnected frequently. This means that relying on these methods without field verification can lead to erroneous decisions, resulting in unlawful killings.

In a well-known case, a Palestinian woman named "Silah" was killed while carrying a white flag and leading her family to safety. After stepping onto a main street, she was targeted by a small drone that shot her in the head. This incident, witnessed by those around

6 Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), Article 57.

7 TRTWORLD (2024) Quadcopter strikes: 1000 Palestinians killed by Israeli drones in one year. Available at Quadcopter strikes: 1000 Palestinians killed with drones in a year.

her, serves as a stark example of the disastrous outcomes of unregulated use of technology on the battlefield[8].

In Libya, drones played a decisive role in the battles between conflicting parties, particularly as many of these drones, including Turkish and Chinese models, were operated using data analysis systems to target objectives. Some of these systems are believed to rely on artificial intelligence techniques for targeting, without legal oversight. The use of these tools in urban areas like Tripoli and Sirte has led to the deaths of civilians and extensive damage to infrastructure[9].

All of these events indicate that integrating artificial intelligence into managing and directing armed conflicts without an internationally binding legal framework to regulate its use could open the door to widespread violations, especially if these systems are not subject to independent and transparent oversight to ensure compliance with international humanitarian law and human rights.

Roles of Companies Developing AI in Armed Conflicts

Through a rapid increase in artificial intelligence and computer services, U.S. tech corporations have discreetly given Israel the ability to monitor and kill many more militants in Gaza and Lebanon more quickly. However, the death toll among civilians has also skyrocketed, raising concerns that these instruments may be causing the deaths of innocent people. Israel's recent wars are a leading example of commercial AI models developed in the United States being used in active warfare, despite concerns that they were not originally designed to help decide who lives and who dies.

For years, militaries have hired private companies to create customized autonomous weapons. Numerous American software companies have backed Israel's battles in recent years, including Microsoft and the San Francisco-based startup OpenAI. Under "Project Nimbus," a $1.2 billion contract signed in 2021[10] when Israel first tried out its in-house AI-powered targeting systems, Google and Amazon offer cloud computing and artificial intelligence services to the Israeli military. The military has made use of Dell and Cisco data centers and server farms. Palantir Technologies, a Microsoft partner in U.S. defense contracts, has a "strategic partnership" that provides AI systems to support Israel's war efforts, while Red Hat, an independent IBM company, has also supplied cloud computing technologies to the Israeli military.

Furthermore, through a number of programs, Microsoft also supplies Israel's government with services that have allegedly been used to help the Israeli military, police, Israeli Prison Service (IPS), and illegal settlement operations. Over 10,000 Palestinians are being held by the IPS as of October 2024; half of them have been detained without being charged or having a trial date scheduled. At least 310 medical professionals, UN employees, women, and children are among the Palestinian prisoners from Gaza who are presently detained in prolonged, secret, and incommunicado detention, where they are subjected to torture, mistreatment, and sexual violence and abuse, according to the UN Human Rights Office.

8 Gaza grandmother gunned down by Israeli sniper as child waved white flag," *Times Kuwait*, November 2024, https://timeskuwait.com/news/gaza-grandmother-gunned-down-by-israeli-sniper-as-child-waved-white-flag.

9 France 24. (2021). "Have Killer Drones Been Deployed in Libya?". France 24. Retrieved from https://rb.gy/1m6k43.

10 APNEWS (2025). As Israel uses US-made AI models in war, concerns arise about tech's role in who lives and who dies. Available at How US tech giants' AI is changing the face of warfare in Gaza and Lebanon | AP News.

Companies are under obligation to respect human rights within their scope of operations. Companies that directly aid the offender – for example, by offering financial, logistical, military, or intelligence support – may be held criminally responsible for a crime committed during an armed conflict. Companies and their managers or executives may be held accountable in certain situations even if they had no direct involvement in the crime or no intention of supporting it. As the Office of the High Commissioner on Human Rights (OHCHR) noted, companies "should treat this risk in the same manner as the risk of involvement in a serious crime, whether or not it is clear that they would be held legally liable[11]."

In light of the concerns raised in this submission and their implications for international peace and security, IHLYI urges states to:

1. **Refrain from the use of AI in military applications**: States should immediately halt the use of artificial intelligence in military activities and establish national regulations and laws to prevent its deployment in warfare.

2. **Work towards a global ban on the military use of AI**: States should actively pursue international agreements and frameworks to ban the use of AI in military contexts, ensuring that no country utilizes AI for warfare.

3. **Avoid the development of autonomous and AI-enabled weapon systems**: States should refrain from developing autonomous weapon systems or AI-powered weaponry that could be used to target humans, ensuring human oversight and decision-making in military actions.

4. **Ensure the protection of personal data**: States must guarantee that personal data is protected from misuse by military forces, law enforcement agencies, border control, and private contractors collaborating with these entities.

5. **Promote accountability in AI development**: Technology companies, researchers, engineers, and financial institutions should commit to not supporting the development or funding of AI technologies designed for military applications, advocating for responsible innovation in line with humanitarian principles.

Peace Movement Aotearoa and Stop Killer Robots Aotearoa New Zealand

[21 May 2024]

Peace Movement Aotearoa and Stop Killer Robots Aotearoa New Zealand welcome the opportunity to contribute our views to the UN Secretary-General's report on artificial intelligence (AI) in the military domain and its implications for international peace and security. Our submission briefly outlines our involvement in this issue, and has three sections summarising our position on: a) A new international instrument on military use of AI and autonomy in weapon systems is urgently needed; b) Key focuses of a new international instrument; and c) Scope of a new international instrument. The points below are based on discussions with our member and supporting groups about the content of this submission.

11 Accessnow (2024) Big Tech and the risk of genocide in Gaza: what are companies doing? https://www.accessnow.org/gaza-genocide-big-tech/.

Introduction

Peace Movement Aotearoa is the national networking peace organisation in Aotearoa New Zealand, established in 1981 and registered as an Incorporated Society in 1982. Our purpose is networking and providing information and resources on peace, humanitarian disarmament, human rights and social issues; and we have extensive national networks of member and supporting groups and individuals. We are a founding member of the Stop Killer Robots campaign and coordinate the national Stop Killer Robots Aotearoa New Zealand (SKRANZ) campaign.

SKRANZ was launched in April 2013 to support the global campaign, with a specific national focus on urging New Zealand to take national action to prohibit the development, production and use of autonomous weapon systems; and to take international action to support negotiations on a new treaty to prohibit autonomy in weapon systems. Since 2023 we have widened our focus to include military use of AI as its perils became increasingly obvious.

(a) A new international instrument on military use of AI and autonomy in weapon systems is urgently needed

As outlined in our submission for the UN Secretary-General's report on autonomous weapon systems (A/RES/78/241) last year, it has been clear for some years now that rapidly developing technological advances in the use of force and increasing autonomy in weapon systems pose an unprecedented threat both to humanity and to the foundations of international human rights and humanitarian law, which are based on respect for human life and dignity, protection of humanity in times of oppression and armed conflict, and human responsibility and accountability for harm.

The serious ethical, humanitarian, legal, and security concerns posed by these developments have been discussed for more than a decade within United Nations bodies – including the Human Rights Council, meetings related to the Convention on Certain Conventional Weapons and in the UN General Assembly – as well as in regional and national governmental and non-governmental forums.

Even as these discussions have taken place, some states have increasingly incorporated autonomy into military use of force in ways that have already resulted in gross violations of international law with disastrous consequences for civilian populations. It is apparent that the absence of specific international law on autonomy in weapon systems, and with differing interpretation by some states as to how existing law applies to new technological developments, the risk of proliferation of ever more dangerous and uncontrollable weapon systems is increasing rapidly.

The need for urgency for international action on this has been highlighted over the past eighteen months by, for example, Israel's use of AI-powered target suggestion systems in Gaza to make high explosive strikes on numerous targets possible in a short time frame, resulting in indiscriminate slaughter of civilians and systematic destruction of life-sustaining infrastructure. The reality of digital dehumanisation with catastrophic consequences is now very evident, as is the increasing tendency towards the development and use of autonomous weapon systems that will remove any remaining vestige of humanity from war.

We have noted with concern that states who brought forward A/RES/79/239 include states that have armed and supported Israel's genocidal attacks on Gaza, and where big data tech companies contributing data storage and AI capabilities to Israel's military systems are based.

Similarly, 'responsible AI in the military domain' (surely an oxymoron) is being promoted by states already developing their own AI targeting and autonomous weapon systems, as a way of undermining the push towards a binding instrument to prohibit these critical threats to international peace and security.

The US 'Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy' has highlighted for us the risk of horizontal proliferation of both military use of AI and autonomous weapon systems as states that do not have their own capability in this regard move from interoperability to integration with the states of armed forces that do or that are developing it. In the case of New Zealand, for example – as it seeks to be a 'combat capable force multiplier with enhanced lethality'[12] – this involves closer military integration with the armed forces of Australia and the US: New Zealand endorsed the US 'Political Declaration' early last year specifically to be compliant with US military doctrine.

These unfortunate developments underscore the urgent need for a new international instrument on military AI and autonomy in weapon systems to clarify and strengthen existing law. The instrument must include both prohibitions and regulations, as outlined below, and must include military use of AI in combat.

As emphasised in the UN Secretary-General's 2024 report on autonomous weapon systems[13], negotiations on a new instrument must begin without any further delay, in a multilateral forum where states can come together to work constructively, where the voices of those whose lives have already been impacted by military use of AI and increasing autonomy in weapon systems can be heard, and where UN agencies, the International Committee of the Red Cross (ICRC), and NGOs are active participants.

(b) Key focuses of a new international instrument

While much of the work around military use of AI and autonomous weapon systems has focused on the issue of meaningful human control over the use of force, it is our view that the key underlying ethical imperative is preventing human beings from being targeted or attacked by any system utilising digital code and/or sensors. A prohibition on military use of AI and autonomy in weapons systems that are designed or used to target human beings must be the starting point.

Meaningful human control over the use of force clearly has an ethical component, but it is also a practical and legal means to ensure accountability for any autonomy in weapon systems that breach the key dictates of humanitarian law.

(c) Scope of a new international instrument

It is our view that a new international instrument should include overarching rules to establish a framework for evaluating current and future technological developments, while promoting increased compliance with international human rights and humanitarian law.

Such overarching rules would prohibit autonomous weapon systems that are designed or used to target humans, and lay out specific obligations to ensure meaningful human control over other systems: for example, that the human operator/s understand the capabilities and limitations of the system, are able to fully evaluate the context in which the system will be used, and are making mindful firing decisions rather than assuming the technology is accurate – this would act to regulate autonomy in weapon systems. It would

12 See, for example, the 2025 Defence Capability Plan released this week.

13 Lethal autonomous weapons systems: Report of the Secretary-General (A/79/88), 1 July 2024.

be useful to specify that decisions made by states on their assessment of new or altered weapon systems that incorporate autonomous features or functions must be transparent.

Furthermore, in the context of the UN Secretary-General's forthcoming report on AI in the military domain and in the light of the awful consequences of military use of AI in Gaza, the scope of a new international instrument must go beyond autonomous weapon systems. It is very clear that there is a spectrum of harmful military use of autonomy, ranging from target decision support systems (as some have described systems such as Lavender), data-based targeting systems, generation of target lists by algorithm or AI, sensor-based targeting systems, through to weapon systems that combine these elements and incorporate varying degrees of machine learning to make target selection decisions and attack autonomously.

We note the 2023 Joint Call by the UN Secretary-General and ICRC President stated *"The autonomous targeting of humans by machines is a moral line that we must not cross"*[14], yet that has already happened – a point reiterated in the UN Secretary-General's 2024 report[15].

It is therefore our view that a new instrument must cover military use of AI – including systems that automate significant decision-making in the use of force, such as target generation, force deployment, and engagement – as well as autonomous weapon systems.

Finally, although we have referred in this submission to military use of AI and autonomy in weapon systems, prohibitions and regulations in a new international instrument must also apply to all coercive agencies of the state, including those used for policing and internal security, for border control, in corrections facilities and in places of detention.

Ploughshares

[11 April 2024]

Project Ploughshares, a Canadian peace research institute, has for over a decade focused its advocacy and research on the military applications of emerging technologies, including artificial intelligence (AI) and autonomous weapons. As AI systems are rapidly advancing and being tested in contemporary conflict zones, international governance frameworks have struggled to keep pace. Meanwhile, intensifying geopolitical competition increases the likelihood that AI technologies will be deployed in complex, dynamic environments for which they are not suited – raising significant risks for civilians.

The wide-ranging use of AI in military applications demands urgent and coordinated international attention. We encourage the Secretary-General and member states to focus on three particularly pressing areas: the use of AI in decision-support systems related to the use of force, the dual-use nature of AI technologies, and the widening capacity gap among states engaging in multilateral discussions.

AI decision-support systems

One area that remains insufficiently addressed in current international discussions is the use of AI in military decision-making, especially decisions about the use of force. Of particular concern are AI-enabled targeting tools such as "Lavender" and "Gospel," reportedly used in in Gaza. These systems are classified as "decision support" because a

14 Joint call by the United Nations Secretary-General and the President of the International Committee of the Red Cross for States to establish new prohibitions and restrictions on Autonomous Weapon Systems, 5 October 2023.

15 As at note 3.

human is technically required to approve target selections. However, there is little transparency regarding how these decisions are made, how frequently AI-generated recommendations are rejected, or whether human operators fully understand how the AI systems reach their conclusions.

In practice, these systems raise the risk of "rubber-stamping," in which human oversight becomes superficial, thereby undermining the principle of meaningful human control and increasing the likelihood of harm to civilians. The potential use of such AI systems in early-warning, surveillance, reconnaissance, and nuclear command-and-control systems further amplifies these concerns.

To mitigate these risks, states must work toward clear norms, regulations, and training requirements that enhance operator understanding, counter automation bias, and ensure genuine human engagement in decision-making processes.

Dual-use challenges

AI's dual-use nature – its applicability to both civilian and military domains – creates further governance complexity. Civilian-developed technologies can be repurposed for military use without appropriate testing or safeguards, increasing the risk of conflict escalation, misuse, and error. Additionally, the accessibility of certain AI tools means that nonstate armed groups may also gain access, potentially using them to target civilians and infrastructure.

We urge states to develop policy mechanisms, including export controls, technology impact assessments, and multistakeholder engagement, to account for dual-use risks and promote responsible innovation.

Capacity- and knowledge-building

Current multilateral discussions reveal stark capacity disparities among states, many of which do not have the resources or technical expertise to participate meaningfully in governance efforts. To ensure inclusive and equitable global engagement, we recommend that states collaborate with the UN Office for Disarmament Affairs to strengthen capacity-building initiatives.

The scientific and academic communities also have a role to play in supporting the development of accessible resources and training materials. International forums, such as the upcoming REAIM Summit in Spain, should include dedicated sessions for knowledge-sharing, especially to support representatives from under-resourced states.

Final thoughts

The international community is at a crossroads. The accelerating militarization of AI demands robust diplomatic responses. We can – and must – move from aspirational principles to concrete, enforceable frameworks, by employing political will, inclusive dialogue, and cross-sector collaboration.

AI-powered warfare is no longer a theoretical risk; it is a present reality. Whether this new era enhances global security or undermines it will depend on the steps states take now to strengthen governance, manage technological competition, and uphold international humanitarian norms.

Without timely, coordinated action, the risks of accidental escalation and unintended conflict will only increase.

Soka Gakkai International

[10 April 2025]

The Soka Gakkai International (SGI) welcomes the opportunity to share our views on the important issue of artificial intelligence (AI) in the military domain. As an NGO whose work is guided by Buddhist principles, we urge that the United Nations, its Member States and other stakeholders take into careful consideration the impact of AI in the military domain from a standpoint of upholding and respecting human dignity.

Introduction

AI in the military domain is rapidly evolving and transforming modern warfare and international peace and security. These systems are being used for various purposes, including surveillance, autonomous weapons, decision-making support, and logistics. With such wide-ranging applications, the integration of AI technologies in military systems poses significant challenges. To better ensure compliance with international humanitarian law (IHL) and uphold protection for civilians and combatants alike there are several issues that we may consider.

Lack of transparency and accountability

- If an AI system were to make an error – such as identifying a target incorrectly – it could be difficult to pinpoint the cause of the error, "the black box problem". Was it a flaw in the data used to train the AI, an issue with the algorithm, or a problem in the operational context? Without transparency within these systems, assigning responsibility is difficult.
- International laws and treaties, such as the Geneva Conventions, were created before AI systems became commonplace in warfare. Without global norms and legal frameworks, there is no consistent approach to ensuring accountability for AI decisions made in warfare.
- With inadequate accountability mechanisms in place, AI could be used for military strategies that violate human rights, suppress civil liberties, or engage in unethical operations.

Speed of decision-making and risk of escalation

- The ability of a military force to make decisions and execute actions faster than its opponent is increasingly viewed as having a strategic advantage. However, the drive for speed can lead to unintended and costly consequences.
- Decisions made too quickly without proper analysis or consideration can lead to poor outcomes, including tactical blunders, strategic missteps, or ethical violations.
- Instead of diffusing a tense situation or negotiating, if combatants react too quickly it could provoke an even greater confrontation, further escalation and prolonged conflict resulting in more human suffering including amongst civilians.
- The acceleration of decision-making processes closes down the possibility of meaningful human control, the growing trend to automate decision-making threatens the ability to achieve human oversight which is essential to facilitate compliance with IHL.

Bias in AI in the military domain

- AI bias refers to the presence of systematic and unfair discrimination in AI systems, such as historical bias, where systems may reinforce harmful stereotypes, bias in data processing and algorithm development which can lead to making biased decisions and bias in how the systems are used.
- AI bias in the military domain is a significant concern, particularly as AI systems are increasingly being integrated into defense and security operations. The potential for AI bias to emerge in these areas can result in human rights implications, exacerbating existing inequalities and lead to deadly consequences for certain groups.
- AI heavily relies on vast amounts of high-quality and reliable data for decision-making. There are several potential violations when it comes to obtaining this data including issues around privacy and surveillance, challenges of bias also arise when dealing with incomplete and inaccurate data.
- When AI systems are biased, they not only perpetuate inequalities but also contribute to the digital dehumanization[16] of marginalized groups.

Proliferation

- Nations may rush to develop AI-based military technologies to outpace their adversaries, which could lead to a destabilizing arms race and increased global tensions.
- Without regulation autonomous weapons systems in particular, could proliferate globally, including amongst non-state actors which could increase crime nationally and regionally, exacerbating social inequalities, overwhelm resources and infrastructures of countries, as well as undermine social and national security.

Conclusion

The issue of AI within the military contexts is complex, and without regulation, it could lead to serious consequences for global peace and security. The desire to speed up decision-making processes within this context has yet to be proven as an effective way of resolving conflicts and achieving peace and security. Furthermore, you cannot divorce AI in the military and AI in civil uses, a failure to address AI in a military context could have widespread repercussions in all spheres of civil life including law enforcement, border control, education, housing and health care. Fundamentally, AI is here to stay, how we utilize it in the military and in our lives will shape the course of humanity. We have the possibility and the responsibility to decide how we want to use technology, knowledge, and the world's resources. To use it in a way that uplifts humanity or degrades it? This is an urgent question that requires moral, ethical and courageous leadership.

Stop Killer Robots

[11 April 2025]

The Stop Killer Robots campaign welcomes the opportunity to submit our views to the United Nations Secretary-General in response to Resolution A/RES/79/239.

16 Digital dehumanization is a process where humans are reduced to data, which is then used to make decisions and/or take actions that negatively affects their lives.

Established in 2012, we are a coalition of more than 270 non-governmental organisations working across 70 countries.[1] We seek to counter threats to humanity and human dignity through the adoption of a new international treaty to prohibit and regulate autonomous weapons systems.[2] We support the development of legal and other norms that ensure meaningful human control over the use of force, counter digital dehumanisation, and reduce automated harm.[3]

Building an effective international response to emerging technologies

Autonomous weapons systems, 'AI in the military domain,' and trends and developments in increasingly automated decision-making and action in the use of force – as well as in our lives and societies more broadly – are all part of the same concerning picture:

The growing influence of computer processing and algorithmic thinking increasingly shapes our interactions in the world and the outcomes available to us. There are clear threats to peace, justice, dignity, human rights, equality, responsibility and accountability, and respect for law. We are getting closer to machine processes determining whom to kill.

To address these challenges effectively, a comprehensive and holistic response is needed from the international community.

Adopting a legally binding instrument on autonomous weapons systems will be one critical component: we must draw basic red lines for humanity against the automation of killing, which brings under jeopardy both international humanitarian law and international human rights law, in particular the presumption of innocence, the right to equality and non-discrimination, dignity, and wipes away contextual circumstances of the target(s) in question. The UN Secretary-General's comprehensive report last year reiterated his urgent call on states to negotiate a legally binding instrument to prohibit and regulate these systems by 2026.

But, a new international treaty on autonomous weapons systems alone may not be enough. States must also reach agreement on preventing and addressing grave harm from other uses of emerging technologies. A whole set of strong international rules are needed that stop the erosion of meaningful human control and the slide towards greater digital dehumanisation and automated harm, across international and domestic practice, in armed conflict and in civilian life.

'Military applications of AI' are already contributing to civilian harm

The risks of integrating AI into the use of force in armed conflict reach far beyond those to peace and security between states: a holistic consideration of peace and security that considers dimensions such as ethical, legal, and humanitarian issues must be taken into account in the UN Secretary-General's report under resolution 79/239.

We are already seeing grave threats to civilian protection and human rights and huge harm being caused by AI and automation in the use of force. This is arising from the quest for speed in warfare, the reduction of people to objects, and issues such as automation bias that Stop Killer Robots has raised the alarm about for years.

1 See www.stopkillerrobots.org/about-us and www.stopkillerorobts.org/a-global-push/member-organisations.

2 See https://www.stopkillerrobots.org/our-policies/.

3 See www.stopkillerrobots.org/vision-and-values/.

We have been horrified by reports of the use of AI-powered 'decision support systems' by Israel in Gaza, which suggest human targets to strike.[4] According to reports, human approval of these suggestions in vast volumes at high speed has been minimal – entailing digital dehumanisation, the erosion of meaningful human decision-making and control (including through automation bias), and directly contributing to massive and devastating harm to civilians in Gaza, alongside other tools.[5]

Many other states are developing and using such 'decision support systems', which raise concerns around international humanitarian law, human rights law, and ethics. So far there are few reports on how these are being deployed, with what constraints and with what impacts. Nevertheless, the push by many states to develop and integrate AI and autonomy into decision-making and the use of force is a huge concern. The further use in hostilities of these kinds of tools by any state in the unacceptable ways that we have seen in Gaza must be prevented. Stop Killer Robots struggles to see how such uses could meet the definition of the responsible application of AI in the military domain given in resolution 79/239.

Further risks to peace and security, rights, and human dignity

The quest for greater speed through AI and automation – towards the goal of increasing the tempo of conflict to a point beyond human cognition in the pursuit of a military and strategic edge – is an extremely dangerous one for international peace and security. These risks are further to the impact 'AI in the military domain' is already having on civilian protection. Risks include unwanted escalation, lowered political thresholds to the use of force, and arms race dynamics.

Technologies that can contribute to target selection (such as threat detection tools) and remote biometric surveillance (such as facial recognition) have already had documented negative impacts on human rights such as the rights to privacy, equality and non-discrimination, freedom of expression and peaceful assembly, and the freedom of movement. In the case of facial recognition for identification (1:n), the technology is considered by many legal experts as wholly incompatible with international human rights law.

That AI systems inevitably encode and reproduce the biases of our societies – including racism, sexism and ableism – and that such bias cannot be eliminated, is also well established. The use of such systems to process people in the use of force will inevitably lead to disproportionate – and multiplied – impacts on already marginalised and minoritised people. Integrating automation and AI into decisions and actions in the use of force against people contributes to digital dehumanisation – the process where humans are reduced to data, which is then used to make decisions and/or take actions that negatively affects their lives.

The relationship with autonomous weapons systems

Stop Killer Robots notes that the UN Secretary-General's report will be on the "application of artificial intelligence in the military domain, with specific focus on areas other than lethal autonomous weapons systems." It is important nevertheless to highlight

4 'Lavender': The AI machine directing Israel's bombing spree in Gaza, +972 Magazine https://www.972mag.com/lavender-ai-israeli-army-gaza/.

5 Questions and Answers: Israeli Military's Use of Digital Tools in Gaza, Human Rights Watch, https://www.hrw.org/news/2024/09/10/questions-and-answers-israeli-militarys-use-digital-tools-gaza.

that various applications beyond the boundary of autonomous weapons systems are closely linked to them.

Firstly, such tools could be integrated as components of autonomous weapons systems now or in the future. For example, a 'decision support system' could be used as an autonomous targeting system, connected to a platform tasked to strike targets on the list generated, based on processing sensor data. Secondly, these tools are linked not only practically, but raise and are part of the same picture of concern. Strikes undertaken based on the nominal human approval of targets generated by a decision support system do not sit far from strikes undertaken with an autonomous weapons system.

It is therefore important that states consider these issues in dialogue: many of the rules and principles developed for autonomous weapons systems on keeping control and rejecting automated killing will need to be extended (with adaptations) to other tools; and, how the development of AI in the military domain more broadly will impact the direction and challenges posed by autonomous weapons systems will need consideration.

Recommendations

Technologies incorporating AI and automation into the use of force in armed conflict are currently being deployed without specific agreed rules; the principles various states have proposed and committed to so far have been too weak and vague to prevent civilian harm and risks to peace and security.

All developments in autonomy and AI in the use of force which threaten our safety, security, and humanity must be urgently and adequately addressed through strong regulation by the international community, with unacceptable uses prevented.

States must:

- Move with urgency to negotiate and adopt a new international treaty to prohibit and regulate autonomous weapons systems;
- In International discussions, critically and meaningfully engage with the implications and real-world consequences of current practice in the use of tools that fall under 'AI in the military domain,' including acknowledging and examining humanitarian harm;
- Fully consider the legal, ethical, humanitarian, and peace and security risks of further development and use of such systems, whatever the perceived 'benefits' may be
- Work urgently to prevent unacceptable uses of technology and trends in development, through committing to develop strong norms for meaningful human control and against digital dehumanisation:
  - This should take place domestically, regionally, and internationally.
  - It must involve a comprehensive and holistic international response, including a legally binding instrument prohibiting and regulating autonomous weapons systems alongside other measures.
  - It should include consideration and development of the other legal instruments necessary to preserve meaningful human control and to protect human dignity against AI in the use of force.

Stop Killer Robots Youth Network

[10 April 2025]

The Stop Killer Robots Youth Network welcomes the opportunity to submit recommendations for consideration by the United Nations Secretary-General in response to Resolution 79/239 "Artificial intelligence in the military domain and its implications for international peace and security" adopted by the General Assembly on 24 December 2024. As a global network of young people under age 30 in over 50 countries working to secure a future free of automated killing, we have consistently advocated for the creation of a new treaty on autonomous weapons systems (AWS) – in particular, we insist on a total prohibition of anti-personnel autonomous weapons as we wish to build a world without such dehumanising weapons. While youth will inevitably face the risks of new weapons technologies, we remain underrepresented in the decision-making process and are often sidelined in forums that shape our interests. As youth who have grown up in an increasingly digital world, we wish to create a future where technology is used to promote peace, justice, equality, and human rights, not perpetuate violence.

With escalating conflicts and the rapid deployment of new weapons technologies around the world, there is an urgent need to reinvest in international law as a measure to build trust and achieve sustainable peace and security. The application of artificial intelligence (AI) in the military domain presents numerous challenges that concern us as youth, including digital dehumanisation, the gamification of violence, and the further erosion of human control and involvement over the use of force.

Military AI & AI systems already in use

Artificial intelligence has been progressively implemented in the military domain over the past decade, however, due to the opacity of military activities and development, the wide public has not been aware of this issue until recently when the active uses of AI systems have been mediatized. We have seen and monitored the use of AI systems to support the targeting of both objects and people. Unfortunately, the use of such systems have not been able to alleviate civilian suffering, for example, in Gaza where one third of victims are children and where too many civilian infrastructures, including critical infrastructures such as humanitarian camps, hospitals[1], and schools[2], have been either directly targeted or indirectly impacted by the hostilities.

There have been other concerning uses[3] of AI systems outside of the military which need to be considered as they might be implemented in the military domain, mainly predictive AI and facial recognition. Predictive AI technologies have been used in the police and judicial systems since the early 2010s and have been shown to be ineffective, incorrect, and subject to reinforcing discriminatory behavior.[4] If predictive AI were to be implemented in the military domain, it could lead to the increasing risk of civilians being targeted as they could be labeled as possible fighters or being indirect victims of military activities due to

1 World Health Organization (2025), 'oPt Emergency Situation Update'. https://www.emro.who.int/images/stories/Sitrep_57.pdf.

2 Save the Children (2025), 'Education Under Attack In Gaza, With Nearly 90% Of School Buildings Damaged Or Destroyed'. https://www.savethechildren.net/blog/education-under-attack-gaza-nearly-90-school-buildings-damaged-or-destroyed.

3 Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica (2016), 'Machine Bias'. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

4 Will Douglas Heaven, MIT Technology Review (2020), 'Predictive policing algorithms are racist. They need to be dismantled'. https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/.

the multiplications of targets with predicted military advantages. Facial recognition technologies (FRTs) are also of concern as they are also unreliable especially when it comes to the identification of non-white males. Facial recognition-enabled targeting in military operations must be prohibited as those systems cannot comprehensively analyse every factor that makes military personnel or civilians a target or not.

Digital dehumanisation

One of the main concerns we have about the use of AI systems in the military domain is the proliferation and banalisation of **"Digital dehumanisation"**. We define digital dehumanisation as the process whereby humans are reduced to data, which is then used to make decisions and/or take actions that negatively affect their lives. This process deprives people of dignity, demeans individuals' humanity, and removes or replaces human involvement or responsibility through the use of automated decision-making in technology.[5] Additionally, the increased speed and scale of target production through military AI erodes moral restraints in war and lowers the impact and capacity of decisions from human operators[6], thus enabling the AI systems to make decisions without meaningful human control, which further dehumanises the decision-making process.

Relying on (Big) data leads to problems

We also believe that the use of (big) data in the military leads to multiple issues which need to be considered.

One of the primary issues is the challenge of data labeling – the process of categorizing and tagging data to train algorithms. Inaccurate or biased labeling can have far-reaching consequences, particularly in the context of distinguishing between combatants and non-combatants in conflict zones. If the data used to train military AI systems is flawed or biased, it can lead to disastrous mistakes, such as the targeting of innocent civilians or misidentification of threats.

A critical issue when relying on big data is that the nature data itself is often broken and is incomplete. This means that the data used to train AI models can be incomplete, outdated, or unrepresentative of real-world situations. Such flaws in data can lead to systems that fail to generalize properly, resulting in inaccurate or incorrect predictions and decisions. For example, in combat situations, a lack of diversity in the data used to identify individuals could lead to inaccurate targeting, with devastating consequences. Important data might be missing or poorly represented, such as the exact location of civilians or combatants, which can lead to AI failing to make informed and balanced decisions. In a war scenario, a system trained with data from a specific past conflict may not be capable of handling a new, unpredictable situation. For instance, an AI system that has been fed data from one particular type of conflict might struggle to apply that data to a war with entirely different characteristics, resulting in errors in target identification or incorrect decision-making.

Another significant problem is that many AI systems operate as black boxes. This means that while these systems make decisions and predictions based on the data they process, the decision-making process is not transparent or easily understood. In military

5 Automated Decision Research (2022), 'Autonomous weapons and digital dehumanisation'. https://automatedresearch.org/news/report/autonomous-weapons-and-digital-dehumanisation-a-short-explainer-paper/.

6 Marta Bo and Jessica Dorsey, OpinioJuris (2024), 'Symposium on Military AI and the Law of Armed Conflict: The 'Need' for Speed – The Cost of Unregulated AI Decision-Support Systems to Civilians'. http://opiniojuris.org/2024/04/04/symposium-on-military-ai-and-the-law-of-armed-conflict-the-need-for-speed-the-cost-of-unregulated-ai-decision-support-systems-to-civilians/.

scenarios, where the consequences of decisions are extremely serious, the lack of transparency is particularly concerning. If an AI system makes an error, such as wrongly identifying a civilian as a combatant, the absence of clarity about how the system reached that conclusion makes it nearly impossible to understand the origin of the error. This makes accountability difficult, as we cannot determine why the system acted in a particular way. The lack of explanation regarding the decision-making processes of AI also makes it impossible to correct or adjust the system's behavior, potentially perpetuating errors without the ability to fix them effectively.

Linguistic and cultural bias embedded in data which is used to train AI systems can create security vulnerabilities and catastrophically misinterpret communications, behaviors, and intentions across diverse cultural contexts, potentially triggering lethal automated responses to misunderstood signals.[7] These systems risk automating and amplifying existing prejudices at unprecedented scale and speed with life or death consequences in conflict zones where cultural misunderstandings could rapidly escalate into devastating military actions causing dire consequences.

Accountability

The inclusion of AI systems in the command and decision-making chains will indubitably lead to a lack of accountability and liability for those relying on these systems to make decisions. It will create a sense of distance and lack of liability on the consequences of a decision which mean that decisions may be made without specific, consistent and thorough analysis of the lawfulness and humane characters of the decision. Then, if an action taken using AI systems violates IHL, the people involved in the implementation and those involved in the decision-making should be held accountable and the use of an AI system shall never exempt people from their responsibilities.

We recognize that military operations are bound by multiple bodies of law – national law, International Humanitarian Law (IHL) and International Human Rights Law (IHRL) – which need to be respected and implemented in order for operations to be lawful. Unfortunately, rules of engagement and of targeting – and all the exceptions – cannot be fully understood and implemented by AI systems. Concepts like doubt, proportionality, and the balance between humanity and necessity are inherently human judgments that cannot be captured by an algorithm. Machines cannot be trusted to uphold these standards on their own. Therefore, it is critical that AI systems never act in a vacuum and that humans retain oversight and decision-making power at all times.

What the future might look like

While AI theoretically has the potential to enhance precision and efficiency in military operations, its integration into warfare raises significant concerns about the future of global security. Autonomous weapons systems, capable of making life-or-death decisions without human control, introduce ethical dilemmas and risks of unintended consequences. The use of AI in military technology is likely to aggravate the existing arms race, as nations compete to develop increasingly sophisticated AI systems, widening the power gap between technologically advanced countries and those less developed, leaving them vulnerable in terms of military readiness. The deployment of autonomous weapon systems and AI-driven tools makes conflict more unpredictable, scalable, and asymmetric, granting certain nations the ability to unleash devastating technologies that smaller states or non-state actors may not

7 Jimena Sofía Viveros Álvarez, Humanitarian Law & Policy (2024), 'The risks and inefficacies of AI systems in military targeting support'. https://blogs.icrc.org/law-and-policy/2024/09/04/the-risks-and-inefficacies-of-ai-systems-in-military-targeting-support/.

be able to counter. The proliferation of AI in the military sphere also raises the threat of terrorism, as organized actors could easily access advanced AI-powered systems. Moreover, the fast-paced, constantly evolving nature of AI development turns military strategies into a "cat and mouse" game, where advancements are met with equally rapid countermeasures. In light of these challenges, the future of military AI must be handled with extreme caution, emphasizing robust ethical frameworks, international regulations, and stringent human oversight to prevent these technologies from destabilizing global peace.

What we need

We call for the establishment of a meaningful legally binding instrument for the use of AI-driven systems in the military requires comprehensive integration of the technical sector alongside state actors, addressing the urgent need for standardized verification protocols and trust-building mechanisms between nations. Such an instrument should define clear autonomy thresholds that specify permissible levels of independence in target selection and engagement, mandate extensive documentation of algorithmic decision processes and testing methodologies and establish explicit red lines that cannot be crossed including prohibited deployment scenarios, target categories, and operational environments. This framework should apply consistently across developing and developed nations, incorporate independent verification bodies with appropriate technical expertise to conduct regular compliance audits, and establish enforcement mechanisms with meaningful consequences for violations, all while facilitating technical data sharing and research that builds confidence between stakeholders in this domain.

These systems present an unprecedented threat to global security and human rights, and the risks they pose to non-combatants are immense. It is crucial that it implements a robust framework of monitoring, accountability and oversight. Firstly, the states need to be bound by positive obligations to ensure the responsible use of AI in the military domain. Accountability is a fundamental aspect of this framework. We call for comprehensive mechanisms that oversee every stage of the AI system life cycle, from development and updates to transfers and research. States must ensure that any uses of AI systems are monitored, with clear reporting structures in place to address incidents promptly. Furthermore, it is vital that human operators using these systems receive thorough training and guidance to make ethical decisions in the field. The principle of meaningful human control must remain central when it comes to the use of AI in the military domain to ensure that ultimate responsibility for any actions remains with human decision makers.

Unione degli Scienziati Per Il Disarmo

[6 April 2025]

Introduction

USPID (***U****nione degli* ***S****cienziati* ***P****er* ***I****l* ***D****isarmo, Union of Scientists for Disarmament*) is an association of concerned scientists – founded in 1983 and based in Italy – which promotes arms control and disarmament initiatives based on scientific understanding of risks posed by military applications of science and technology. USPID submits to the United Nations Secretary-General its views on "Artificial intelligence in the military domain, with specific focus on areas other than lethal autonomous weapons systems, and its implications for international peace and security", in accordance with the invitation formulated in operative paragraphs 7 and 8 of Resolution 79/239 adopted by the UN General Assembly on 24 December 2024.

Hazards for peace and security arising from AI military applications

USPID expresses its deep concern about new hazards for peace, international security, and the respect of International Humanitarian Law (IHL) which arise on account of the ongoing and accelerating military efforts to incorporate Artificial Intelligence (AI) into multiple facets of warfare. Major sources of these hazards have been identified in current limitations of our capability to understand, predict precisely, and control the behavior of AI systems developed by machine learning methods and their interactions with other human or artificial agents. Initially identified in connection with the operation of AI-enabled Autonomous Weapons Systems (AWS), these hazards are now spreading to AI systems supporting intelligence collection, the achievement of situational awareness, and human decision-making in warfare.

Exceptionally grave concerns are raised by proposals to integrate AI in Nuclear Command, Control, and Communication (NC3) and in adjacent systems supporting nuclear decisions, and to let AI perform tasks that might directly or indirectly affect nuclear decision-making. A significant case in point is the proposal to use AI technologies in nuclear early warning and decision-support systems, which is being advanced with the expectation that AI accuracy will reduce potential errors, and its processing speed will buy more time for nuclear decision makers. However, on account of the probabilistic nature of AI information processing, one cannot exclude the risk of AI perception leading to false positives of a nuclear attack or producing perniciously unreliable recommendations given the impossibility of ensuring that the underlying models are aligned with human values and the UN overarching goal of preventing and removing threats to peace. If such mistakes occur, no matter how infrequent, large-scale and even existential implications for humanity might ensue. Accordingly, it would be imperative to proceed with time-consuming verifications of AI responses in nuclear early warning. But these verifications would be hindered by the black-box nature of much AI information processing and by the reliance on mostly simulated data, eventually thwarting the expectation of buying more time for human decision makers.

Additional concerns are raised by proposals to exploit the rapid pace at which AI operates to speed up battlefield decision-making and targeting cycles. These proposals are fueled by the goal of gaining military advantage over potential adversaries. However, fighting at machine speed jeopardizes both the effectiveness of human oversight on AI-enabled decision support systems and the fulfilment of ethical and legal roles that are attached to human oversight of warfare action. Indeed, overly tight temporal windows for decision-making hinder effective human control over IHL threats raised by machine suggestions. Human interventions which aim at preventing inadvertent conflict escalations prompted by fighting at machine speed are similarly hampered. In addition to this, excessive speed in human-machine interactions has been identified as a factor inducing automation biases on the battlefield, and potentially skewing human decision-making even in the absence of AI failures.

Further hazards arise in connection with inherent vulnerabilities of AI learning methods and systems. Malicious manipulation of input data might be exploited to induce classification mistakes by AI systems. Moreover, poisoning attacks corrupting learning datasets may impair learning processes and the accuracy of resulting AI systems. These risks are compounded by our current inability to fully align AI systems with human goals and values, potentially causing them to deviate from strategic objectives.

Recommended actions

Mindful of these and other emerging hazards posed by the rapid adoption of AI technologies and systems in the military domain, USPID recommends

- to integrate discussion of AI in NC3 into the Non-Proliferation Treaty framework and in dedicated high-level dialogues and forums such as the Summit on Responsible Artificial Intelligence in the Military Domain (REAIM);
- to develop sustained international dialogue, good practices, and confidence-building measures concerning new and emerging risks for peace and IHL respect raised by AI warfare applications;
- to support a comprehensive and detailed inquiry aimed at identifying actual and potential AI applications in the military domain, jointly with situations of use that pose serious threats to peace, international stability, and the respect of IHL;
- to consider and investigate the need to introduce international regulations or prohibitions for those AI military applications that pose serious threats to peace, international stability, and the respect of IHL.

Women's International League for Peace and Freedom

[24 May 2024]

The Women's International League for Peace and Freedom (WILPF) has opposed war and the development of technologies of violence since its founding in 1915. WILPF has consistently condemned military spending and militarism as detrimental to human life and wellbeing. Our concerns with artificial intelligence (AI) in the military domain and its implications for international peace and security are grounded within our wider opposition to weapons, war, and violence, as well as in our opposition to patriarchal, racist, and colonial power relations that are embedded within AI technology.

While there are many perils of the military use of AI; WILPF's submission is focused on the following issues:

1. The need for human emotion, analysis, and judgement in relation to the use of force;
2. The existence of gender, racial, and other bias in AI technology and the implications for digital dehumanisation;
3. The impacts of military use of AI on privacy and personal data;
4. The environmental harms exacerbated by the military use of AI; and
5. The dangers of war profiteering and arms racing.

Due to the concerns raised in this submission and in other spaces, WILPF opposes the military use of AI. This technology, rather than placing limits on violence or harm, expands both. Governance is insufficient in the face of the profits and power the developers of these technologies seek.

In light of the concerns raised in WILPF's submission and the implications for international peace and security, WILPF urges states:

- To refrain from using AI in the military domain and to develop national laws and regulations to this end;
- To pursue a global prohibition on the military use of AI;
- To not develop autonomous weapon systems or AI-enabled weapon systems, including those that can be used to target human beings;

- To ensure protection of personal data from use by militaries, police, border enforcement, and private companies and contractors collaborating with these institutions;
- To uphold human rights and dignity online and offline; and
- To address the environmental harms generated by data centres, cloud computing, and AI by reducing the number of these centres and energy consumption and water use, which will include reducing the overall use of AI.

WILPF also urges:

- Technology companies, tech workers, scientists, engineers, academics and others involved in developing AI or robotics to pledge to never contribute to the development of AI technologies for military use;
- Financial institutions such as banks and pension funds to pledge not to invest money in the development or manufacture of AI for military use; and
- Activists, academics, affected communities, and other concerned about privacy rights, digital dehumanisation, environmental and climate justice, gender-based violence, and other issues to collaborate and strategise to oppose the development and use of AI in the military and other violent domains.

D. Scientific Community

AI, Automated Systems, and Resort-to-Force Decision Making Research Project, The Australian National University

[11 April 2025]

Introduction

This executive summary highlights policy recommendations outlined in ***AI, Automated Systems, and Resort-to-Force Decision Making – Policy Recommendations: Submission to the UN Secretary General Pertaining to A/RES/79/239*** **(11 April 2025)**, available on the UNODA website. For a complete account of the underlying research and associated research papers, please refer to the full submission.

Underlying Research Project

This research has arisen from a **two-and-half-year research project (2022-2025),** entitled ***Anticipating the Future of War: AI, Automated Systems, and Resort-to-Force Decision Making***, led by Professor Toni Erskine (Australian National University) and funded by the Australian Government through a grant by the Department of Defence.

Its focus is **distinctive and critical**. While the attention of academics and policy makers has been overwhelmingly directed towards the use of AI-enabled systems in the *conduct of war* – including, prominently, on the emerging reality of 'lethal autonomous weapons systems' ('LAWS'), this project has addressed the **relatively neglected prospect of employing AI-enabled tools at various stages and levels of deliberation over the *resort to war.*** In other words, 'it takes us from AI on the battlefield to **AI in the war-room'**.[1]

1 T. Erskine and S. E. Miller, 'AI and the Decision to Go to War: Future Risks and Opportunities', *Australian Journal of International Affairs*, Vol. 78: 2 (2024), pp. 135–147 (p. 138).

This research project has brought together **leading scholars and practitioners** working on different aspects of international politics and security, strategic and defence studies, and artificial intelligence (AI) to contribute to a multi-disciplinary study and set of **policy recommendations on the risks and opportunities of introducing AI, machine learning (ML), and automated systems** into state-level decision making on the **initiation of war**. Our interventions are made from the perspectives of political science, international relations, law, computer science, philosophy, sociology, psychology, engineering, and mathematics.

Project participants presented and discussed their research at two workshops (June 2023 and July 2024) at the Australian National University (ANU), convened by Professor Toni Erskine and Professor Steven E. Miller (Harvard). Participants also received feedback on their initial research-based policy recommendations from senior Australian Government delegates from the federal civil service as part of a one-day policy roundtable (July 2024) at the ANU.

'Four Complications'

For all the potential **benefits** of AI-driven systems – which are able to analyse vast quantities of data, make recommendations and predictions by uncovering patterns in data that human decision makers cannot perceive, and respond to potential attacks with a speed and efficiency that we could not hope to match – challenges abound. Through this project, we have sought to address **four thematic 'complications'** that we propose will accompany the gradual infiltration of AI-enabled systems in **decisions to wage war:**[2]

- **Complication 1** relates to the displacement of human judgement in AI-driven resort-to-force decision making and possible implications for deterrence theory and the unintended escalation of conflict.
- **Complication 2** highlights detrimental consequences of automation bias, or the tendency to accept without question computer-generated outputs – a tendency that can make human decision makers less likely to use (and maintain) their own expertise and judgement.
- **Complication 3** confronts algorithmic opacity and its potential effects on the democratic and international legitimacy of resort-to-force decisions.
- **Complication 4** addresses the likelihood of AI-enabled systems impacting organisational structures and chains of command, whether degrading or enhancing strategic and operational decision-making processes.

Contributors to this project have explored these proposed complications in the context of either **automated self-defence** or the use of **AI-driven decision-support systems (DDS)** that would inform human resort-to-force deliberations. We have identified risks and opportunities of using AI-enabled systems in these contexts and make recommendations on how risks can be mitigated and opportunities promoted.

Complication 1: Displacement of human judgement

AI in Nuclear Crisis Decision Making

One key area of research undertaken in response to this complication is the nuanced interplay between AI and human decision making in the high-stakes context of **nuclear**

2 For an account of these 'four complications', see T. Erskine and S. E. Miller, 'AI and the Decision to Go to War: Future Risks and Opportunities', *Australian Journal of International Affairs,* Vol. 78: 2 (2024), pp. 135–147 (pp. 139–40).

crisis management. **Risks (including the increased fragility of nuclear deterrence relationships, crisis signalling becoming more complex**, and **unintended escalation)** have been explored in two broad areas: i) automation in military deployments, or taking the human 'out of the loop' in the decision to use nuclear or strategic non-nuclear weapons (SNNW); and, ii) the integration of AI into human decision-making (particularly in early warning threat assessments). Although much of this research has focused on risks, **novel benefits** of introducing AI-driven decision-support systems (DSS) into human-led nuclear crisis management have also been proposed.

Policy Recommendations:

- **Always incorporate human-in-the-loop safeguards**: Ensure AI systems in nuclear command and control are always overseen by human operators and that human decision-making remains central to determining when and how nuclear-weapon states resort to the use of their arsenals.
- **Promote a holistic approach to AI-safety**: AI safety should account for both technical and socio-technical dimensions. Assess safety challenges in AI-enabled DSS comprehensively, including issues of security, trust, and liability.
- **Broaden the scope of risk assessments**: Apply risk assessments relating to the deployment of AI and ML not only to obvious areas such as nuclear launch orders, but also to less obvious areas such as early warning intelligence assessments (including by non-nuclear allies) and SNNW capabilities (including by non-nuclear allies).
- **Restrict the use of AI-assisted warning data**: The key to balancing the benefits of incorporating AI into early warning against the risks is limiting what AI-assisted warning data is used for. In AI research, prioritise tasks such as calculating effective evasive manoeuvres in the event of an attack and using pattern recognition and anomaly detection to improve arms control verification.
- **Pursue** informal arms control and confidence-building: Advance informal measures such as regular dialogue, red-line agreements, and information-sharing mechanisms. Expand unilateral initiatives like moratoriums where feasible.
- **Explore AI's potential to promote empathy and enhance decision making**: Decision makers must exercise 'security dilemma sensibility' (SDS) in times of crisis. Decision makers and diplomats exercise SDS when they are open to the possibility that the other side is behaving the way they are because they are fearful and insecure, and crucially, recognize the role that their own actions may have played in this. Explore ways that the balanced integration of AI and human judgement could enhance SDS during nuclear crises by promoting empathy and trust.

AI Mistakes in the Resort to Force

Another area addressed in relation to this complication is **state responsibility** when **errors** occur in AI-driven or autonomous systems involved in resort-to-force decisions. Such errors may arise from poor system training, data poisoning by adversaries, or two AI-driven systems interacting in unintended ways. It is essential to develop legal standards and practices that reduce the risk of unintended conflict resulting from such failures.

Policy Recommendations:

- **Adopt robust security and cyber hygiene**: States should adopt robust protections against AI data poisoning and cyber attacks to meet jus ad bellum standards of good faith and reasonable conduct.
- **Clarify legal guidelines on delegating the use of force to autonomous systems**: Senior leadership within states should set clear domestic legal standards regarding when and how autonomous systems may be authorised to use force.
- **Commit to transparency in after-action reviews**: States should commit to being transparent and deliberate about after-action reviews of any AI errors that occur in the field, potentially drawing on civilian casualty review processes as a model.

Complication 2: Automation bias

Our research in response to the second complication focuses on the relationship between human actors and **AI-driven DSS** in resort-to-force decision making. It includes a detailed survey-based study of **military trust in AI** during strategic-level deliberations and a robust account of the importance of ensuring that there are human **'experts-in-the-loop'** when AI-driven systems contribute to decisions on war initiation. This body of work also addresses the **benefits** of employing DSS to **enhance our cognitive capacities** in strategic decision making and, conversely, uncovers the potential **dangers** of such reliance if these systems **dull our sensitivity to the tragic qualities of war** or contribute to the **erosion of restraint** by creating the illusion that they replace us as responsible actors.

Policy Recommendations:

- **Consider the multidimensionality of trust**: Recognize that soldiers' trust in AI is not a forgone conclusion. Rather, it is complex and multidimensional, and further complicated by biases, uncertainty, and lack of education.
- **Interrogate norm compliance**: In terms of governance, explain how policies on increasingly autonomous capabilities coincide or diverge from international norms and laws informing their use.
- **Embed experts in decision structures**: Enshrine an 'expert-in-the-loop' organisational structure – i.e., high-level experts as core decision makers.
- **Prohibit automation**: Prohibit automation of resort-to-force decisions.
- **Increase AI literacy of domain experts**: Provide and require basic technical training for high-level domain experts so they understand the logics of AI and can thus incorporate AI decision inputs from an informed position.
- **Provide on-going, substantive training for domain experts**: Sustain substantive training for, and assessment of, high-level experts to bolster and ensure substantive competencies.
- **Regulate non-autonomous AI**: While autonomous AI agents, e.g., lethal autonomous weapons systems (LAWS), need regulation, so do non-autonomous AI systems, which leave humans vulnerable to new forms of influence, moral and cognitive atrophy, and undermined responsibility.
- **Design AI-driven DSS to promote more accurate perceptions of their capacities**: Ensure AI-driven DSS are not easily mistaken for responsible agents in themselves by

avoiding anthropomorphism, building in warnings about system limitations, and incorporating features that emphasise human agency and accountability.

Complication 3: Algorithmic opacity

Our research in response to the third complication addresses **how the lack of transparency of AI-driven decision making can threaten the legitimacy** of AI-informed decisions on the resort to force. This body of work includes original research on **large language models (LLMs)** and their potential to exacerbate existing **pathologies in intelligence analyses**. It also examines the role that the **'architecture of AI'** and its hidden vulnerabilities play in deliberations surrounding the resort to force. Moreover, research within this pillar conceives of military decision-making institutions as **'complex adaptive systems'** – a conceptual framework that yields a range of insights, including that human-machine teams possess a form of **'cognitive diversity' that could be leveraged for more efficient decision-making,** but also **exploited to poison information flows**, and that technical explanations for algorithmic opacity will not solve accountability concerns.

Policy recommendations:

- **Develop policy to limit epistemic pathologies of LLMs**: States should clearly determine defence and intelligence policy towards either a) procurement of LLMs, b) state development of LLMs, or c) a combination of both. They should use this guidance to develop policy which seeks to limit the epistemic pathologies of LLMs in autonomous decision-making.
- **Commit to sector-wide procurement guidelines and oversight of generative AI tools** used in decision-making chains.
- **Commit to regulating data markets** and access to those markets through alliance relationships.
- **Promote understanding of the tech ecosystem and its fragilities**: Increase understanding of the inherent interdependencies and vulnerabilities in the tech ecosystem, including by creating technology literacy training programs designed specifically for politicians and policy, intelligence, and military leaders.
- **Invest in research** to develop a comprehensive picture of the architecture – physical and digital – that underpins AI, including critical dependencies and vulnerabilities and how access and power are distributed.
- **Invest in research on social media** and its impact on functions of government, including its potential to disrupt democracies, facilitate foreign interference, and influence decision making on the use of force.
- **Recognize AI's current influence**: Significantly increase awareness of government reliance on the architecture of AI, especially for critical government functions, including resort-to-force decision making.
- **Invest in research and development** to maximize the benefits of human-machine cognitive diversity.
- **Implement responsible AI governance programs** that carefully balance accountability with operational efficiency.
- **Perform regular red-team exercises** to ensure that the integration of AI in decision-making institutions does not induce systemic blind spots and vulnerabilities in military decision-making.

Complication 4: Impact on organisational structures

Our research regarding the fourth complication explores both the **beneficial and damaging effects that AI-driven systems can have on institutional structures** in the context of resort-to-force decision making. Studies focus on how AI-driven **DSS can improve 'adaptive culture'** within military organisations, thereby improving wartime decisions, and how the urgent need to **upgrade AI literacy and educate human analysts** should lead us to **reform institutional structures and cultures**. The novel notion of **'proxy responsibility'** is proposed as an institutional response to ensure that responsibility can be meaningfully assigned to humans for resort-to-force decisions that are informed by AI systems. Moreover, original research highlights the significance of the **neglected category of AI 'integrators'** – sandwiched between the 'developers' and 'users' of AI within organisational structures – when it comes to strategic military applications of AI.

Policy recommendations

- **Set (and evolve) measures of effectiveness**. If AI-enabled adaptive capacity is to work effectively, measures of military effectiveness must guide which direction adaptation might take. Establish such measures at the tactical (battlefield) and strategic (war-room) levels to guide development and implementation of AI-enabled adaptation.
- **Know where adaptation relevant data is found, stored and shared**. An enhanced adaptive stance in military institutions must have enhanced data awareness as a foundation. Data awareness and management must become one of the basic disciplines taught to military personnel.
- **Scale AI support from individual to institution**. There is unlikely to be a one-size-fits-all algorithm or process that can enhance learning and adaptation at every level of military endeavours. Create a virtual 'arms room' of adaptation support algorithms as part of an institution-wide approach to adaptation.
- **Routinely question AI-enabled outputs**: Build mindsets, protocols, institutional cultures, and inter-agency structures in 'normal' pre-crisis times to routinely question AI-enabled output from human-machine teams.
- **Institute an advisory body**: In order to support the notion of 'proxy responsibility' as an institutional response to 'responsibility gaps' when decisions on war initiation are informed by AI-enabled systems, establish and/or strengthen state-level 'AI departments'. These departments would integrate technical, political, and ethical competence and expertise and advise on resort-to-force decision-making processes.
- **Support research on AI integration**: Fund research on the integration of AI in strategic decision-making.
- **Provide standards**: Outline minimum standards for the responsibilities of AI developers and integrators.
- **Facilitate inter-group discussions** between developers, integrators and users during development, integration, and longer-term maintenance processes.
- **Create accountability guidelines**: Provide well-defined guidelines and rules indicating who is accountable if something goes 'wrong'.

Queen Mary University of London, T.M.C Asser Institute, University of Southern Denmark, University of Utrecht

[11 April 2025]

Views of members of the scientific community and civil society; specifically, we are a group of academics with expertise in ethical, legal and political dimensions of military Artificial Intelligence and herewith put forward our shared views pursuant to resolution 79/239 "Artificial intelligence in the military domain and its implications for international peace and security" adopted by the General Assembly on 24 December 2024, in accordance with the request of the UN Secretary-General contained in Note Verbale ODA/2025-00029/AIMD.

Introduction:

The rapid advancement and integration of AI technologies into targeting operations have sparked ongoing debates surrounding their ethical, legal, and operational implications. Over the past decade, the discourse on AI in warfare has largely centered on autonomous weapon systems (AWS),[1] driven in part by the initiation of discussions in 2013 and the formalization of a regulatory process under the UN Convention on Certain Conventional Weapons (CCW) and the Group of Governmental Experts on Lethal Autonomous Weapons Systems (GGE LAWS), which exclusively focuses on lethal AWS.[2] However, the increasing integration of AI-based decision-support systems (AI-DSS) into targeting practices[3] introduces new layers of complexity that demand closer attention from a broad range of stakeholders. This submission responds to that need, structured around three key components: (1) a brief overview of how AI-DSS are currently used in targeting decisions; (2) an analysis of key concerns, including how these systems shape the potential exercise of human judgement and control and underline fundamental gaps in global governance; and (3) a concluding set of recommendations.

1. Overview of AI-DSS and the joint targeting cycle

Defined as "the process of selecting and prioritizing targets and matching the appropriate response to them, considering operational requirements and capabilities,"[4] targeting is a core military function at the very heart of warfare. While the potential range

1 The latest definition of AWS from the CCW GGE LAWS Rolling Text (26 November 2024): "A lethal autonomous weapon system can be characterized as an integrated combination of one or more weapons and technological components that enable the system to identify and/or select, and engage a target, without intervention by a human user in the execution of these tasks." On file with authors.

2 For a brief overview of some of the latest developments of the GGE LAWS see Jeroen van den Boogaard, *Warning! Obstacles Ahead! The Regulation of Autonomous Weapons Systems in the GGE LAWS*, *Opinio* Juris, 4 March 2024 found at: https://opiniojuris.org/2024/03/04/warning-obstacles-ahead-the-regulation-of-autonomous-weapons-systems-in-the-gge-laws/.

3 There have been several reported uses of AI-DSS by Israel in Gaza and potentially in Lebanon, by both Ukraine and Russia in the ongoing conflict, and by the United States in its actions against Houthi rebels in the Red Sea and in Yemen, to name a few. For a comprehensive overview of literature in this space, see e.g., Anna Nadibaidze, Ingvild Bode, and Qiaochu Zhang, *AI in Military Decision Support Systems, A Review of Developments and Debates,* Centre for War Studies, University of Southern Denmark, November 2024. Found here: https://www.autonorms.eu/ai-in-military-decision-support-systems-a-review-of-developments-and-debates/.

4 United States Department of Defense, *Dictionary of Military and Associated Terms*, March 2017, found at: https://www.tradoc.army.mil/wp-content/uploads/2020/10/AD1029823-DOD-Dictionary-of-Military-and-Associated-Terms-2017.pdf.

of use cases for AI-DSS in military decision-making is broad, in targeting, AI-DSS can be understood to serve as **tools** that use AI techniques to collect and analyze data, provide information about the operational environment as well as actionable recommendations, with the aim of aiding military decision makers in evaluating factors relevant to legal compliance such as taking precautions and ensuring proportionality in attacks.

More specifically, AI-DSS are increasingly integrated across multiple phases of the joint targeting cycle (JTC), including within target development and prioritization, capabilities analysis, and mission execution. The JTC is a reflective example of a structured process used by military forces to identify, evaluate, and engage targets while ensuring compliance with operational, legal, and ethical standards,[5] generally consisting of six (non-linear) phases:

1. **End-State and Commander's Objectives**: Defining strategic military goals and desired outcomes.

2. **Target Development and Prioritization**: Identifying, verifying/validating, and prioritizing targets based on intelligence and mission goals.

3. **Capabilities Analysis**: Assessing the available strike options and their effectiveness.

4. **Force Assignment**: Allocating specific military assets (e.g., airstrikes, artillery, cyber operations) to engage the target.

5. **Mission Execution**: Carrying out the targeting operation while ensuring compliance with relevant laws and the rules of engagement.

6. **Assessment**: Evaluating the effectiveness of the operation and adjusting for future operations, if necessary.

Within this framework, AI DSS are assumed to serve primarily as informational and analytical tools which support human decision-making rather than supplant it. However, this assumption and framing obscures how AI-DSS influence human cognitive processes within the JTC. This impact on human decision-making is often underestimated and remains insufficiently examined, leaving critical discussions about the role of AI-DSS largely absent from current policy debates.

2. Analysis of Key Concerns

(a) (Meaningful) Human Judgement and Control

AI-DSS are often portrayed as enhancing human decision-making and the quality of decisions therein. The perception of AI-DSS as mere subsidiary tools has led to a narrative that the integration of AI-DSS poses fewer challenges than AWS, given that these systems do not directly "engage" targets (i.e., they do not have an inherent capability to directly carry out the use of force) and are tools that assist human commanders. The outputs are ostensibly

5 Michael Schmitt et al, *Joint and Combined Targeting: Structure and Process*, Chapter 13 in Jens David Ohlin (ed) *Weighing Lives in War* (Oxford, 2017). See also, Jessica Dorsey and Marta Bo, *AI-Enabled Decision-Support Systems in the Joint Targeting Cycle: Legal Challenges, Risks, and the Human(e) Dimension*, forthcoming 2025, *International Law Studies.* "Targeting generally involves four key steps: (1) objectives and guidance, (2) planning, (3) execution, and (4) assessment. Encapsulating these four key steps, the United States and NATO outline their targeting processes through similar six-phase cycles [addressed in this submission]. As the reader can discern, different states employ different doctrines for targeting. What is important … is not necessarily the specific labels for various steps followed by any given state, but rather how and when compliance with the principle[s of IHL are] incorporated into the targeting process."

ultimately reviewed through (several layers of) human oversight, such as processes of verifying and validating targets using additional intelligence sources.[6] As a result, errors or inaccuracies in AI-DSS outputs are often seen as non-critical, based on the assumption that robust human oversight and appropriate control will compensate for them. However, closer examination reveals that this control is frequently superficial, offering only the appearance of, rather than actual meaningful, or context-appropriate, human judgement and control.

This is because AI-DSS structure and condition the quality of human control and oversight and limit the ways control and oversight can be exercised. The use of AI-DSS creates a shared decision-making space between human military personnel and AI technologies. States appear to have recognized and focused on many of the advantages of this shared decision-making space for military personnel, i.e., how the use of AI-DSS advances human decision-making through offering data-driven insights. But using AI-DSS also delimits the capacity to exercise human oversight and control because of the technologies' complexity and the increased speed (and therefore scale) it can bring to decision-making processes. Rather than supporting human oversight, using AI-DSS may risk humans becoming little more than reactive cogs in socio-technical systems.[7] Moreover, this configuration risks amplifying adverse human biases, such as automation bias, anchoring bias, or cognitive action bias, to the detriment of exercising qualitatively high levels of human control.[8] Considering AI-DSS as a distinct form of technology therefore reveals significant challenges associated with military AI and human oversight, challenges that extend beyond those that arise when simply integrating the technology in weapon systems.

Recent conflicts have shown the risks associated with AI-DSS being employed in critical functions, such as target selection and even nomination, and their conditioning and constraining of human involvement, affecting the fulfilment of core legal obligations embedded within the JTC. The use of AI-DSS raises fundamental concerns about whether human decision makers can retain adequate cognitive autonomy over the JTC process or whether humans will become overly reliant on algorithmic outputs for critical judgements in the context of armed conflict.[9] Consequently, there are significant legal concerns regarding the effects of such systems on decision-making processes and use of force decisions and ability for users to comply with IHL obligations, especially with respect to the obligation to take all feasible precautions to minimize civilian harm to the greatest extent possible in attack and comply with the principles of distinction and proportionality.[10]

Importantly, these concerns are not new. There is extensive debate around how to preserve meaningful human judgment and human agency when conducting IHL-evaluative legal assessments, in the context of AWS. These discussions – which include expert analysis on accountability, human-machine interaction, automation bias, and the effect of AI systems

6 Alexander Blanchard and Laura Bruun, *Automating Military Targeting: A Comparison Between Autonomous Weapon Systems and AI-Enabled Decision Support Systems*, Stockholm International Peace Research Institution (SIPRI) forthcoming 2025 (draft on file with authors).

7 Ingvild Bode, *Human-Machine Interaction and Human Agency in the Military Domain*, Policy Brief No. 193 (Waterloo, ON: Centre for International Governance Innovation, 2025).

8 Dorsey and Bo, supra n. 5.

9 *Ibid*; see also Anna Nadibaidze, Ingvild Bode, and Qiaochu Zhang, *AI in Military Decision Support Systems, A Review of Developments and Debates,* Centre for War Studies, University of Southern Denmark, November 2024. Found at: https://www.autonorms.eu/ai-in-military-decision-support-systems-a-review-of-developments-and-debates/.

10 Article 57 of the First Additional Protocol to the Geneva Conventions. See also Dorsey, Bo supra n. 5 (on AI-DSS and their effects on the principle of precautions); Jessica Dorsey, *Proportionality under Pressure: The Effects of AI-Enabled Decision Support Systems, the Reasonable Commander Standard and Human(e) Judgment in Targeting*, forthcoming *International Review of the Red Cross* (2025) (on AI-DSS and their effects in the context of IHL proportionality assessments).

on legal and ethical reasoning[11] – provide valuable lessons that must inform discussions around military AI and specifically the use of AI-DSS in military contexts.

(b) AI-DSS: Understudied, Under-Addressed and Unregulated

Framing AI-DSS as mere tools, has led to an underestimation and lack of analysis on the way their use affects the cognitive decision-making process within the JTC. The relative lack of attention paid to AI-DSS so far can partly be attributed to the fact that such systems are seen to be used with a human *in* or *on* the loop framework, with their outputs ostensibly reviewed by one or more individuals during the targeting process. As a result, current understandings of AI-DSS use appear to align with widely supported principles of human control and oversight. However, this gap in the debate is also caused by a lack of transparency around how specific AI-DSS function, and a consistent failure to comprehensively examine how they are being used in practice.

Additionally, the persistent focus on AWS at the expense of AI-DSS obscures the growing reliance on AI in shaping operational and strategic outcomes. Unlike AWS, which have been debated in the framework of the CCW for the past decade, AI-DSS lack a comparable institutional platform. Attention to AI-DSS remains scattered across various initiatives but these efforts have yet to provide the dedicated regulatory focus or coordination needed.

3. Recommendations:

i. **Reassert** the central role of human cognitive and legal reasoning in military operations by implementing safeguards that ensure key legal assessments remain grounded in human(e) judgment. Leverage existing insights from debates on AWS and research on human-machine teaming and human-computer interaction to inform discussions on AI-DSS.

ii. **Recognize** and address the incremental effects of AI-DSS design and use on human cognitive reasoning and critical deliberation. Promote awareness and attentiveness as a crucial part of reasserting and strengthening the exercise of human agency in targeting decision-making.

iii. **Reinforce** calls for greater attention to the implications of AI-DSS in armed conflict. Utilize platforms such as the UN General Assembly's First Committee on Disarmament and International Security and the Global Commission on the

11 Marta Bo, *Autonomous Weapons and the Responsibility Gap in light of the Mens Rea of the War Crime of Attacking Civilians in the ICC Statute*, 19 Journal of International Criminal Justice 2021; Bo, M., Bruun, L. and Boulanin, V., *Retaining Human Responsibility in the Development and Use of Autonomous Weapon Systems: On Accountability for Violations of International Humanitarian Law Involving AWS* (SIPRI: Stockholm, Oct. 2022), p. 41; Boulanin, V., Bruun, L. and Goussac, N., *Autonomous Weapon Systems and International Humanitarian Law: Identifying Limits and the Required Type and Degree of Human-Machine Interaction* (SIPRI: Stockholm, June 2021), p. 54; and Bruun, L., Bo, M. and Goussac, N., *Compliance with International Humanitarian Law in the Development and Use of Autonomous Weapon Systems: What Does IHL Permit, Prohibit and Require?* (SIPRI: Stockholm, Mar. 2023), p. 24. Elke Schwarz, "The (im)possibility of meaningful human control for lethal autonomous weapons systems," *Humanitarian Law and Policy*, 29 August 2018, found at: https://blogs.icrc.org/law-and-policy/2018/08/29/im-possibility-meaningful-human-control-lethal-autonomous-weapon-systems/.

Responsible Use of AI in the Military Domain to foster inclusive and complementary discussions on the associated risks and systemic changes AI-DSS introduce.

United Nations Institute for Disarmament Research

[11 April 2025]

Artificial intelligence (AI) is rapidly transforming the military domain and profoundly influencing international peace and security. Initiatives such as the summits on Responsible AI in the Military Domain (REAIM) and the Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy, while not being universal processes, have significantly elevated international attention on the military applications of AI. In particular, they have moved the debate beyond lethal autonomous weapon systems (LAWS) and have successfully highlighted the multifaceted impacts of AI, fostering broader international policy engagement. Building on the political momentum generated by these initiatives, resolution 79/239 adopted by the United Nations General Assembly in December 2024 represented a significant milestone as the first UN resolution on AI in the military context and has offered Member States, international and regional organizations and the multi-stakeholder community the opportunity to share their views on opportunities and risks.

For many years, the United Nations Institute for Disarmament Research (UNIDIR) has played an important role in shaping and informing discussions on the broader impact of AI in the military domain, both within and beyond applications of this technology in weapon systems. It has undertaken research, facilitated multilateral dialogues, and offered policy insights that underline AI's transformative potential for international peace and security. This policy note draws from all the work conducted to summarise opportunities and risks and to offer a potential roadmap for future policy action.

The international community can now shape how AI is used in the military domain, putting principles of responsible AI at the core. A central challenge is the complexity of defining the "military domain". States and regions interpret the scope of this domain differently based on their unique security landscapes, realities and operational practices. For some countries, military roles extend to internal security tasks such as policing, border control, combating organized crime, protection of critical infrastructure or humanitarian relief in response to natural disasters. Others maintain a stricter definition, limiting military functions to battlefield engagements. These variations, rather than serving as barriers, offer important context for multilateral discussions. International governance frameworks should remain flexible and inclusive, acknowledging and adapting to diverse national and regional security perspectives.

In the many operational contexts within the military domain, AI acts as a force multiplier across several military tasks, including command and control (C2), information management and intelligence, advanced autonomy, logistics, training and simulation, and organizational and support functions. In C2, AI enhances the speed and quality of decision-making, thereby helping commanders rapidly analyse battlefield scenarios. It has the potential to improve adherence to international humanitarian law (IHL), for example by integrating detailed proportionality and other legal assessments. AI-driven intelligence tools analyse large volumes of data at speed, and so improve situational awareness and threat detection. In logistics, AI optimizes supply chains and predictive maintenance, enhancing operational readiness and improving the sustainability of military operations over time. AI further supports advanced autonomy in drones, cybersecurity, and operations in the information domain. Training and simulation benefit from AI by creating personalized, realistic synthetic environments and scenarios. In short, if developed, deployed and used responsibly, AI could increase operational effectiveness while offering new ways to mitigate risks and reduce harm.

However, integrating AI in military contexts also presents significant risks and challenges – technological, security, legal, policy and ethical.

Technologically, military AI systems face issues related to the quality, availability and inherent biases of data. These may lead to unpredictable and potentially harmful outcomes, including violations of international law. The "black box" nature of AI systems, often coupled with their adaptiveness and highly context-dependent nature, complicates trustworthiness assessments and may, at times, challenge the conduct of effective investigations into alleged violations of IHL. Cybersecurity vulnerabilities also expose AI systems to adversarial attacks, requiring stringent security measures.

Security challenges include risks of miscalculation and unintended escalation, particularly through AI-enabled rapid decision-making processes and AI-enabled autonomy, which may result in escalatory responses. The potential for an AI arms race exacerbates international and regional tensions, possibly leading to destabilizing outcomes similar to historical arms competitions. The proliferation of AI technologies to non-state actors further complicates threat landscapes and necessitates robust life cycle management of military AI systems. Additionally, AI-generated disinformation threatens societal stability by undermining trust in information and can have a direct impact on military operations.

Legal challenges revolve around ensuring compliance with international law, particularly IHL and international human rights law. Key debates focus on, among other things, accountability and both state and individual responsibility for AI-driven actions, especially regarding lethal decisions. States diverge on whether existing legal frameworks are sufficient or if new, specialized regulations are required. Beyond international law, ethical considerations emphasize maintaining human judgment in critical decision-making and preventing societal biases from infiltrating AI systems. The latter requirement calls for greater diversity and inclusivity in AI development. Additionally, bridging gaps between government, academia and the private sector remains challenging yet crucial for effective governance.

Addressing these challenges requires a comprehensive road map with actions at the multilateral, regional and national levels.

Multilaterally, establishing a United Nations-led comprehensive platform that enables a regular institutional dialogue to address military AI's broader implications on international peace and security is key as it would provide an institutional framework to advance policy discussions. This platform could build on the existing internationally developed AI principles and frameworks, such as UNESCO's recommendations or the commitments made in the Global Digital Compact (e.g. safe, secure and trustworthy AI) and further refine them for application in the military domain. These principles could be further developed into voluntary norms of responsible behaviour in the development, deployment and use of AI in the military domain and provide a solid foundation for future multilateral instruments. In addition, such platform could be leveraged to develop practical confidence-building measures (CBMs), lead inclusive multi-stakeholder engagement, and deliver global capacity-building programmes that enhance global security via transparency, cooperation and predictability.

Regionally, existing organizational frameworks can be used to tailor CBMs and guidelines that reflect local security contexts. Cross-regional dialogues would facilitate mutual learning, prevent information silos, and include diverse perspective which would encourage globally coherent responses.

Nationally, states should develop comprehensive AI strategies that detail vision, priorities and governance frameworks, ensuring compliance with international norms and ethical standards. Robust governance structures (e.g., dedicated AI steering committees and ethics boards), alongside iterative legal reviews, would enhance accountability and safety. Transparent communication and clearly defined accountability protocols would further support responsible AI implementation. High standards of data governance, life cycle management approaches, rigorous training programmes and updated military operational

guidelines complete these proposed national measures, ensuring the responsible integration of AI in the military domain.

Table A below provides an overview of the proposed roadmap for policy action.

Table A: A roadmap for future policy action

Level	*Action*	*Rationale*
Multilateral	**Establish a multilateral process under United Nations auspices to provide a comprehensive platform for discussion on military applications of AI and their impact on international peace and security.** This process could be leveraged to: a. Develop a set of overarching, core principles of responsible AI in the military domain to help align national efforts and reduce risk. b. In the future, further develop these core principles into international voluntary norms or guidelines for responsible state behaviour in the development, deployment and use of AI in the military domain. These guidelines could take the form of a code of conduct or a political declaration supplemented by more technical instruments as required (e.g., on AI assurances, and robust protocols for testing and evaluation). c. Develop confidence-building measures (CBMs) for military AI. States could agree on and implement practical CBMs to increase transparency and trust regarding AI in the military domain. d. Promote multi-stakeholder engagement in support of multilateral policy action. e. Develop and implement a coherent capacity-building programme.	Collectively, these multilateral actions aim to foster cooperation, set common rules and share knowledge on military AI at the international level with a view to increasing predictability. They aim to shape the global landscape so that all states move towards safer and more transparent integration of AI in the military domain, thereby reducing the risks. While clustered under a single umbrella recommendation, each of the actions above could be implemented on its own, although their mutually reinforcing nature would amplify the impact achieved if they are implemented in combination.
Regional	**Leverage regional and subregional organizations and dialogues to discuss the issue of AI in the military domain.** Regional and sub-regional organizations could: a. Develop region-specific CBMs, norms or guidelines that reflect local contexts.	Regional and subregional approaches allow tailoring to specific security realities and threat perceptions, which could lead to concrete results that are more aligned with specific needs. In addition, regional and subregional approaches could be leveraged to inform and shape global dialogues and strengthen context-specific capacity-building.

Level	*Action*	*Rationale*
	b. Set up networks for information-sharing on AI-related best practices suited to their security landscape. c. Develop joint AI-development projects, aligning operational, legal and technical requirements. **Initiate cross-regional dialogues Initiate cross-regional dialogues on AI**, where two or more regional groups exchange lessons and possibly align their approaches.	Cross-regional dialogue can be a useful tool to enable mutual learning and avoid echo chambers.
National	**Implement a comprehensive approach to AI governance in the military domain** to include the following actions: a. Develop a comprehensive national strategy or policy on AI in security and defence. b. Establish robust governance structures and review processes. c. Implement transparency and accountability measures d. Implement robust data practices and governance frameworks for all military AI applications. e. Manage AI capabilities throughout their entire life cycle – from design and development, through testing and deployment, to updates and decommissioning – with continuous risk assessments and mitigation at each stage. f. Invest in human capital and training by developing extensive training programmes for military personnel on AI and cultivating a new generation of AI-literate officers and specialists. This includes not only technical training but also training on the ethical and legal aspects of AI use in operations. g. Review military operational guidelines to strengthen AI governance in military contexts, including military documentation (e.g. doctrines, standard operating procedures and others), and rules of engagement.	A national strategy clarifies roles and responsibilities, and provides a clear direction for the development, acquisition, integration and use of AI in the military domain. Dedicated structures provide focus and accountability. They create effective checkpoints that AI projects must pass and comply with consistently (e.g., ethical approval, legal clearance, safety certification), reducing chances of unsafe or unlawful deployment. Transparency builds public trust and international confidence that a state is using AI responsibly. Accountability ensures that the presence of AI does not create a vacuum of responsibility – maintaining the ethical and legal norm that humans are accountable for military actions. By prioritizing robust data governance and the provision of the necessary infrastructure to enable it, militaries can improve the performance and trustworthiness of their AI systems and reduce error rates. A life cycle view ensures that safety and compliance are ongoing commitments reducing chances of failure in the field and ensuring that accountability is maintained throughout the system's use. Human expertise and judgment remain critical. Training reduces misuse and enables more effective human-machine teaming. Existing military governance tools and instruments can be used to strengthen the

Level	*Action*	*Rationale*
		governance of AI in the military domain at a more practical, tactical level, thereby offering an impactful complement to the highest levels of governance and the associated obligations emanating from international, regional and national laws and regulations.

E. Industry

Microsoft

[24 May 2024]

Microsoft welcomes the opportunity provided by the United Nations General Assembly resolution A/RES/79/239 on "Artificial Intelligence in the Military Domain and its Implications for International Peace and Security", and UNODA's invitation to share perspectives on the opportunities and challenges posed to international peace and security by the application of artificial intelligence (AI) in the military domain, with specific focus on areas other than lethal autonomous weapons systems.

Our perspectives reflect Microsoft's deep commitment to our Responsible AI Principles and our Secure Future Initiative, emphasizing cybersecurity, safeguarding international norms, and promoting trust in technology, and our active participation in multi-stakeholder initiatives including the UNIDIR-led Roundtable for AI, Security, and Ethics (RAISE).

I. Opportunities

Microsoft recognizes substantial opportunities in responsibly applied AI within the military domain, particularly:

- *Enhancing cybersecurity and defense capabilities:* AI significantly strengthens cybersecurity defenses by automating threat detection, enabling faster and more accurate responses to cyber threats. Technologies such as Microsoft Security Copilot illustrate the transformative potential of AI in defense, empowering cybersecurity professionals to identify and mitigate risks efficiently. Initiatives like Microsoft's Zero Day Quest and collaboration with MITRE ATT&CK demonstrate proactive industry efforts to enhance global cybersecurity preparedness and resilience.
- *Broad spectrum of military applications:* Beyond cybersecurity, responsibly designed AI can significantly enhance efficiency and effectiveness across logistics, command and control systems, intelligence processing, military training, peacekeeping, humanitarian assistance, and disaster relief operations. Diverse applications underscore AI's transformative potential beyond combat scenarios alone.
- *Improving compliance with international humanitarian law:* AI technologies should improve the accuracy and effectiveness of targeting processes, aiding militaries to better adhere to principles of distinction, proportionality, and necessity. AI should significantly enhance protections for civilians and civilian infrastructure, thereby reducing unintended collateral damage in conflict.

- *Capacity building and international cooperation:* The adoption of AI in the military domain presents opportunities for global knowledge-sharing and capacity-building initiatives. International partnerships should support developing nations by sharing security capabilities, knowledge, and best practices, thus bridging technological divides and fostering global stability.

II. Challenges

Microsoft equally acknowledges significant challenges and risks associated with AI applications in the military domain:

- *AI-enhanced cyber threats:* AI has escalated cyber threat capabilities, empowering state-sponsored and criminal actors to carry out increasingly sophisticated cyber operations. These AI-driven threats include advanced phishing campaigns, automated exploitation of vulnerabilities, and adaptive malware, significantly increasing global cybersecurity risks.
- *Risks of escalation and miscalculation:* Integrating AI into military decision-making risks unintended escalation and/or miscalculation. Rapid, automated decision-making processes may inadvertently lower conflict thresholds, amplifying risks of destabilization or accidental conflict.
- *Proliferation and uncontrolled diffusion:* Uncontrolled diffusion, especially through open-source models and decentralized development, heightens the risk of malicious use by both state and non-state actors, including terrorist groups and cyber mercenaries. Increasingly accessible dual-use and proprietary AI systems enable actors even with limited resources can gain access to capabilities that previously required significant investment or expertise, posing additional threats to international security and stability.
- *Algorithmic bias and ethical implications:* Algorithmic biases embedded within AI systems pose ethical and humanitarian concerns. Biases related to gender, race, age, or socioeconomic factors in AI datasets can intentionally and unintentionally perpetuate inequality and discrimination, particularly within sensitive military and security applications.
- *Digital divides and inequality:* Without deliberate policy actions, disparities between developed and developing nations in AI capabilities could deepen, increasing geopolitical tensions and socio-economic inequalities, thus undermining long-term global stability.

III. Relevant normative proposals

Microsoft recognizes several existing and emerging normative frameworks relevant to AI governance in the military domain, including:

- UNIDIR's RAISE initiative, facilitating international multi-stakeholder dialogues and governance proposals.
- The Responsible AI in the Military Domain (REAIM) Summits, emphasizing transparency, accountability, and human oversight at the international level.
- The US Department of Defense Responsible AI Strategy, highlighting responsibility, equitability, traceability, reliability, and governability.
- NATO's Principles of Responsible Use for AI in Defence, emphasizing reliability, governability, and traceability among member nations.

IV. Microsoft recommendations

To maximize opportunities and mitigate the challenges, Microsoft proposes several key recommendations:

- *Establish clear international norms and standards:* Develop explicit international norms and industry standards governing responsible use and development of military AI. These norms should delineate acceptable and unacceptable behaviors, providing robust frameworks to deter misuse and foster transparency and accountability, supported where appropriate by monitoring or compliance mechanisms. AI governance frameworks should explicitly differentiate operational contexts, such as peacekeeping, humanitarian assistance, crisis management, and conflict scenarios, to appropriately address varied ethical, legal, and humanitarian considerations. To ensure continued relevance, such norms should be periodically reviewed and updated to reflect evolving technological developments and operational realities.

- *Ensure human-centric oversight and accountability:* Adopt policies ensuring meaningful human judgment, oversight, and accountability remain central to military decisions involving AI, particularly regarding the use of force. Clear oversight mechanisms and enforceable accountability structures, including rigorous human control and review processes, are necessary to maintain ethical standards, avoid automation bias, and mitigate unintended consequences.

- *Advance secure and transparent AI development practices:* Promote rigorous technical standards and comprehensive life cycle management protocols covering pre-design, development, testing, deployment, operation, acquisition, and decommissioning. Robust vulnerability management, security audits, and transparent development and deployment processes should be integral components, alongside clear capacity-building measures, ensuring AI systems remain secure, responsible, and resilient throughout their operational life cycle.

- *Enhance responsible data governance practices:* Establish clear international guidelines on responsible data governance specifically tailored to military AI applications. Transparent and accountable data management practices addressing collection, sharing, storage, training, and operational usage are crucial for managing dual-use risks, preventing misuse, and maintaining strict compliance with international legal and ethical frameworks.

- *Address and reduce algorithmic bias:* Prioritize addressing algorithmic bias through rigorous testing, transparent data practices, and inclusive AI development processes. Developers and users should establish clear policies to proactively identify, mitigate, and remediate biases, especially when AI systems are deployed in sensitive military or security contexts.

- *Promote responsible innovation and risk-based regulation:* Support regulatory frameworks that are risk-based, outcome-oriented, and balanced, ensuring they encourage innovation while adequately addressing security and ethical risks associated with AI deployment. Industry should advocate for flexible, adaptive regulations that keep pace with technological change, without imposing overly prescriptive or impractical requirements. Industry-led initiatives, such as voluntary codes of conduct, vulnerability disclosure standards, and collaborative red-teaming exercises, should be actively supported and integrated into broader international normative frameworks.

- *Strengthen international governance and alignment:* Support and actively engage in international initiatives, including REAIM Summits and dialogues at the UN General Assembly and UN Security Council. Robust international governance frameworks,

characterized by transparency, clear accountability measures, and trust-building mechanisms, are essential for coherent and inclusive approaches to AI governance. Member States and stakeholders should coordinate closely through these forums to reduce fragmentation and ensure global alignment.

- *Support knowledge-sharing and awareness-raising with the UN system:* Encourage and actively contribute to efforts by the UN Secretariat and relevant UN entities to convene meaningful multi-stakeholder expert dialogues, workshops, and knowledge-sharing on AI in the military domain. Exchanges through voluntary contributions, technical expertise, and collaborative initiatives should aim at enhancing global understanding of AI's implications for international peace and security.
- *Strengthen international cooperation and information sharing:* Encourage robust international cooperation, emphasizing real-time threat intelligence sharing and joint attribution mechanisms. Industry actors should actively participate in collective cybersecurity efforts, enhancing global cybersecurity preparedness and response.
- *Foster multi-stakeholder dialogue and collaboration:* Actively participate in and support forums such as RAISE, involving states, international organizations, academia, civil society, and industry. Such inclusive dialogues are essential for mutual understanding, shaping responsible AI practices, and developing collaborative governance structures.

V. Conclusion

Microsoft is deeply committed to proactive collaboration with Member States, the UN system, industry, and civil society to implement these recommendations swiftly and effectively. Through sustained collective efforts and ongoing engagement in multi-stakeholder initiatives, Microsoft will continue supporting responsible AI governance, innovation, and practices that meaningfully contribute to international peace and security.