

An ninh con người và công nghệ quân sự mới nổi

Không có gì
quá thông minh?

Ứng dụng quân sự của
trí tuệ nhân tạo

**DRONE
WARS**

Giới thiệu



Hình minh họa của Quân đội Hoa Kỳ về lực lượng mặt đất đang tuần tra bằng máy bay không người lái và hệ thống tự động. Tín dụng: Quân đội Hoa Kỳ

Trí tuệ nhân tạo (AI), ra quyết định tự động và công nghệ tự động đã trở nên phổ biến trong cuộc sống hàng ngày và mang đến những cơ hội to lớn để cải thiện xã hội một cách đáng kể. Điện thoại thông minh, công cụ tìm kiếm trên internet, trợ lý cá nhân AI và ô tô tự lái là một trong nhiều sản phẩm và dịch vụ dựa vào AI để hoạt động. Tuy nhiên, giống như tất cả các công nghệ, AI cũng gây ra rủi ro nếu nó không được hiểu rõ, không được kiểm soát hoặc được sử dụng theo những cách không phù hợp hoặc nguy hiểm. Cùng với việc biến đổi ngôi nhà và doanh nghiệp, AI được các cường quốc quân sự thế giới coi là một cách để cách mạng hóa chiến tranh và giành lợi thế trước kẻ thù. Các ứng dụng quân sự của AI như công cụ lập kế hoạch hậu cần DART (xem bên dưới) đã được sử dụng hàng ngày trong vài thập kỷ qua và các hệ thống mới với những đặc điểm đáng lo ngại đang nhanh chóng được triển khai.

Bản tóm tắt này là một trong loạt bài do Drone Wars UK xuất bản như một phần của dự án 'Chiến tranh tương lai' của chúng tôi. Đây là phiên bản rút gọn của một báo cáo dài hơn, cả hai đều xem xét các ứng dụng quân sự của AI và mô tả cách quân đội Vương quốc Anh bắt đầu áp dụng các công nghệ AI, trước khi tiếp tục phác thảo những rủi ro khác nhau liên quan đến chúng. Phiên bản đầy đủ có sẵn tại www.drone wars.net/None-too-clever-full.

AI là gì?

Các nghiên cứu học thuật thường thừa nhận rằng không có định nghĩa thống nhất chung về AI vì sự đa dạng của các phương pháp tiếp cận nghiên cứu trong lĩnh vực này. Chính phủ Vương quốc Anh đã mô tả AI là “Các công nghệ có khả năng thực hiện các nhiệm vụ đòi hỏi trí thông minh của con người, chẳng hạn như nhận thức trực quan, nhận dạng giọng nói và dịch ngôn ngữ”.¹ Thảo luận về các phương pháp tiếp cận khoa học máy tính và kỹ thuật khác nhau được định nghĩa này đề cập đến là ngoài phạm vi của phần tóm tắt này, mặc dù điều quan trọng là phải chỉ ra rằng các hệ thống AI thường có khả năng học hỏi hoặc thích ứng với thông tin hoặc tác nhân kích thích mới.²

AI dựa vào việc xử lý lượng lớn dữ liệu để xác định các mẫu thống kê. Bất chấp nhiều thập kỷ nghiên cứu, cho đến gần đây vẫn chưa có đủ năng lực máy tính cũ ng như bộ dữ liệu mở rộng cần thiết để cho phép các hệ thống AI hoạt động. Sự gia tăng theo cấp số nhân về tốc độ xử lý và dung lượng lưu trữ của máy tính, tốc độ internet và 'điện toán đám mây', cùng với số hóa hàng loạt và thu thập lượng dữ liệu khổng lồ, hiện đang cho phép AI trở nên trư ờng thành.

AI thường đư ợc phân loại là 'hẹp' hoặc 'chung'.

Các ứng dụng hiện tại đều là ví dụ về AI thu hẹp, trong đó máy móc thực hiện một nhiệm vụ cụ thể cho một mục đích cụ thể.

Thuật ngữ chung “các phương pháp tính toán” có lẽ là một

cách tốt hơn để mô tả những hệ thống như vậy, vốn kém xa trí thông minh của con người như ng có khả năng giải quyết vấn đề có thể áp dụng tổng quát hơn so với phần mềm thông thường. Khi hệ thống AI hoạt động bằng cách nhận dạng các mối quan hệ thống kê trong các tập dữ liệu, chúng thường đư ợc sử dụng cho các chức năng sau:

- Tự động hóa các nhiệm vụ.
- Xử lý các tập dữ liệu phức tạp hoặc lớn.
- Dự đoán hành vi.
- Gắn cờ những điều bất thường hoặc sự kiện quan tâm.
- Gắn thẻ dữ liệu và sửa lỗi.³

AI tổng quát là khả năng giả thuyết của một hệ thống máy tính để thực hiện một loạt các chức năng nhận thức và phản hồi nhiều loại dữ liệu đầu vào cũ ng như hiểu và giải quyết bất kỳ vấn đề nào mà bộ não con người có thể làm đư ợc. Mặc dù đây là mục tiêu của một số chương trình nghiên cứu AI như ng nó vẫn là một viễn cảnh xa vời.⁴

AI không hoạt động biệt lập mà đóng vai trò là “xử ơng sống” trong một hệ thống rộng hơn để giúp hệ thống đạt đư ợc mục đích của mình. Người dùng không 'mua' chính AI; họ mua các sản phẩm và dịch vụ sử dụng AI hoặc nâng cấp hệ thống cũ bằng công nghệ AI mới. Hệ thống tự trị là những cỗ máy có thể thực hiện nhiệm vụ mà không cần con người

đầu vào, dựa vào hệ thống máy tính trí tuệ nhân tạo

¹ 'Chiến lược công nghiệp: Xây dựng một nước Anh phù hợp cho tương lai'. Vụ Chiến lược Kinh doanh, Năng lực và Công nghiệp, tháng 11 năm 2017. P37. https://www.gov.uk/Government/uploads/system/uploads/attachment_data/file/664563/công-nghiệp-chiến-lược-trắng-giấy-web-ready-version.pdf

² Michael Copeland: 'Sự khác biệt giữa Trí tuệ nhân tạo, Học máy và Học sâu là gì?' Blog của Nvidia, ngày 29 tháng 7 năm 2016. <https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/> 3

Lawrence Lewis: 'AI và quyền tự chủ trong chiến tranh: Hiểu và giảm thiểu rủi ro'. CNA Analysis & Solutions, tháng 8 năm 2018. P6. https://www.cna.org/CNA_tập_tin/PDF/DOP-2018-U-018296-Final.pdf

⁴ Lawrence Lewis: 'AI và quyền tự chủ trong chiến tranh: Hiểu và giảm thiểu rủi ro', op cit. P8.



Lễ cầu nguyện tưởng niệm 290 nạn nhân dân sự trên chuyến bay 655 của Iran Airlines bị USS Vincennes bắn hạ năm 1988. Hệ thống phòng không AEGIS của tàu khiến các nhà khai thác có rất ít thời gian để giải thích dữ liệu mà nó cung cấp. tin dụng Raheleh Zomorodinia

5 Ingvald Bode và Tom Watts: 'Kiểm soát con người ít ý nghĩa hơn: Bài học từ Hệ thống phòng không về kiểm soát con người có ý nghĩa đối với cuộc tranh luận về AWS'. Drone Wars Vũ khí quốc Anh và Đại học Nam Đan Mạch, tháng 2 năm 2021. <https://dronewars.net/wp-content/uploads/2021/02/DW-Control-WEB.pdf>

để diễn giải thông tin từ các cảm biến và sau đó truyền tín hiệu cho các bộ truyền động, chẳng hạn như động cơ, máy bơm hoặc vũ khí, để vận hành một cơ chế gây tác động đến môi trường xung quanh máy.

Các hệ thống tự trị có thể hoạt động thông qua những gì đôi khi được mô tả là tự chủ khi nghỉ ngơi và tự chủ khi chuyển động.

Quyền tự chủ ở trạng thái nghỉ mô tả các hệ thống hoạt động trong phần mềm hoặc trong thế giới ảo, trong khi quyền tự chủ khi chuyển động mô tả các hệ thống tương tác với thế giới vật lý.

thế giới. Trong bối cảnh quân sự, cả hai loại hệ thống đều có thể gây lo ngại. Các hệ thống tự chủ ở trạng thái nghỉ có thể đưa ra các quyết định quan trọng về việc sử dụng vũ lực và có thể gây ra hậu quả lớn, trong khi các hệ thống tự động chuyển động có thể bao gồm các hệ thống vũ khí tự động gây chết người (LAWS) - 'robot sát thủ' có thể tự đưa ra quyết định về mục tiêu và giết chóc mà không cần sự can thiệp của con người.

Việc sử dụng tự động hóa và điện toán trong các hệ thống quân sự không phải là mới. Hệ thống vũ khí tự động đã tồn tại từ Thế chiến thứ hai, với những tiến bộ trong công nghệ máy tính dần dần góp phần tạo nên sự tinh vi của chúng.

Vào những năm 1940, một số radar máy bay và phòng không đã được đưa ra trang bị bộ tiếp sóng mà chính người điều khiển radar và hệ thống radar có thể xác định xem máy bay mà họ đang theo dõi là thân thiện hay thù địch. Khi công nghệ tiến bộ, các hệ thống phòng không có thể xác định máy bay và tên lửa bằng cách so sánh tốc độ, thông tin radar và dấu hiệu đầu của chúng với cơ sở dữ liệu, đồng thời có thể tự động điều chỉnh hướng đi và xác định mục tiêu bằng radar hoặc cảm biến tầm nhiệt. Các hệ thống phòng không hiện đại như hệ thống tên lửa Patriot hoặc Aegis, được thiết kế để hoạt động chống lại nhiều mối đe dọa đang đến với tốc độ cao, có thể đưa ra quyết định nhắm mục tiêu để con người phê duyệt hoặc thậm chí tấn công mục tiêu mà không cần sự giám sát của con người. Điều này đã được chứng minh là có vấn đề sâu sắc: trong những hệ thống như vậy, người điều khiển con người bị giảm vai trò đóng một vai trò tối thiểu như ng vô cùng phức tạp, và các hệ thống phòng không tự động đã dính vào một số lỗi khiến máy bay dân sự và quân sự bị bắn rơi trong các sự cố hỏa hoạn giao hữu. .5

Quân đội và AI

Doanh nghiệp và học viện đã dẫn đầu và tiếp tục dẫn đầu sự phát triển của AI vì họ có vị trí tốt hơn để đầu tư vốn và tiếp cận các nguồn lực cần thiết cho nghiên cứu so với quân đội và khu vực công. Do đó, rất có thể các ứng dụng quân sự trong tương lai của AI sẽ là sự thích ứng của các công nghệ được phát triển trong lĩnh vực thương mại.

Các tổ chức nghiên cứu của chính phủ như Cơ quan Dự án Nghiên cứu Tiên tiến Quốc phòng (DARPA) và Cơ quan Dự án Nghiên cứu Tiên tiến Tình báo (IARPA) ở Hoa Kỳ và Phòng thí nghiệm Khoa học và Công nghệ Quốc phòng (DSTL) ở Anh có nhiều dự án AI nhằm mục đích khuyến khích hợp tác với các lĩnh vực thương mại và học thuật để thích ứng và sử dụng AI và các công nghệ tự hành cho mục đích quân sự.

Mặc dù cuộc họp này chủ yếu dựa trên các ví dụ từ Mỹ và Anh, nhưng Trung Quốc, Nga và các cường quốc quân sự khác cũng có các chương trình tích cực để phát triển AI quân sự.

Đặc điểm của AI khiến nó trở nên hấp dẫn đối với quân đội. Các tính năng như tốc độ nhanh, khả năng xử lý các tập dữ liệu lớn và phức tạp để tìm ra các mẫu và thực hiện các nhiệm vụ lặp đi lặp lại một cách chính xác đã định hình vai trò của nó được các lực lượng vũ trang giao phó. Bảng 1 cho thấy một số thuộc tính của AI được quân đội đặc biệt quan tâm.

Bảng 1: Các chức năng hỗ trợ sử dụng quân sự của AI6

Tốc độ phân tích và hành động nhanh chóng.
Thực hiện các nhiệm vụ tự động đơn giản ở quy mô lớn.
Điều khiển hệ thống robot và tự động.
Nhận biết các mẫu để dự đoán xu hướng trong tương lai hoặc phát hiện sự bất thường.
Phân loại và nhận dạng đối tượng và tín hiệu.
Tối ưu hóa hệ thống để đạt được mục tiêu.
Nâng cao chất lượng của việc ra quyết định.

Các phần tiếp theo xem xét một số lĩnh vực mà AI đang nổi lên như một công nghệ quân sự. Trong một số ứng dụng này, AI đang cho phép hoặc có thể cho phép phát triển các sản phẩm hoặc hệ thống mới trong tương lai. Trong các trường hợp khác, AI có thể được sử dụng để giúp họ hoạt động hiệu quả hơn, kéo dài tuổi thọ hoặc giao cho họ những vai trò mới hoặc được cải thiện. Máy bay do thám U-2 của Không quân Hoa Kỳ, bay lần đầu tiên vào năm 1955, gần đây đã được điều khiển bởi một phi công phụ AI có khả năng học hỏi và thích nghi, không giống như chế độ lái tự động thông thường. Phi công phụ AI, có nhãn hiệu 'Artup' (Artoo), là một thực thể dựa trên thuật toán, trong chuyến bay huấn luyện, có thể điều khiển các cảm biến của máy bay và truyền thông tin về vị trí của các bộ phóng tên lửa cho phi công con người. Các phi công phụ AI dự kiến sẽ thường xuyên bay cùng các phi công của Lực lượng Không quân Hoa Kỳ là con người trong tương lai gần và Artup được cho là có cơ hội tốt được sử dụng trong các hoạt động vào mùa hè năm 2021.7

6 Dựa trên Forrest E. Morgan, Benjamin Boudreaux, Andrew J. Lohn, Mark Ashby, Christian Curriden, Kelly Klima, Derek Grossman: 'Ứng dụng quân sự của trí tuệ nhân tạo: Mối lo ngại về đạo đức trong một thế giới không chắc chắn'. Tập đoàn RAND, 2020. https://www.rand.org/pubs/research_reports/RR3139-1.html

7 Valerie Insinna: 'Phòng thí nghiệm công nghệ nhỏ bé đưa AI lên máy bay do thám còn có một dự án bí mật khác'. Tin tức Quốc phòng, ngày 11 tháng 2 năm 2021. <https://www.defensenews.com/air/2021/02/11/the-tiny-tech-lab-that-put-ai-on-a-spyplane-has-another-secret-project/>

Ứng dụng quân sự của AI

Làm việc từ xa

Các hệ thống robot và tự động theo truyền thống được sử dụng để thực hiện các hoạt động quân sự 'buồn tẻ, bẩn thỉu hoặc nguy hiểm' nhằm thay thế con người điều khiển. Công nghệ tự động có thể ngăn chặn việc con người phải bước vào những tình huống nguy hiểm trong môi trường nguy hiểm. Ví dụ bao gồm phát hiện và xử lý chất nổ, hoạt động quét mìn trên biển hoặc trên đất liền hoặc trinh sát trong môi trường thù địch.

Tình báo, Giám sát và trinh sát

Việc thu thập thông tin tình báo tập hợp các tập hợp dữ liệu lớn, bao gồm tài liệu văn bản, video và hình ảnh tĩnh, thông tin chặn điện tử và thông tin nguồn mở từ internet. AI đang bắt đầu đóng vai trò xử lý và phân tích dữ liệu này.

Một ví dụ về việc sử dụng AI trong phân tích tình báo là Dự án Maven - chính thức được gọi là Nhóm chức năng chéo chiến tranh thuật toán (AWCFT) - được Lầu Năm Góc thành lập vào tháng 4 năm 2017 để phát triển việc sử dụng AI để phân tích dữ liệu. Vào cuối năm đó, các thuật toán AI được phát triển thông qua Project Maven đã được sử dụng tại hiện trường để hỗ trợ phân tích các cảnh quay video bằng máy bay không người lái ở Trung Đông.

Mặc dù Project Maven, theo như được biết, tập trung vào việc sử dụng AI để nhận dạng đối tượng, hệ thống AI có thể được sử dụng để hỗ trợ một loạt hoạt động của con người ở các giai đoạn khác nhau trong quy trình phân tích trí thông minh. Ở giai đoạn đầu của quy trình, AI có thể được sử dụng để lọc và phân loại tài liệu được thu thập hàng loạt, chẳng hạn như bằng cách sử dụng thuật toán nhận dạng giọng nói hoặc tín hiệu để 'dọn dẹp' dữ liệu khỏi môi trường ồn ào. Ở giai đoạn tiếp theo, hệ thống AI có thể hỗ trợ phân tích dữ liệu thô, ví dụ như thông qua dịch máy và tóm tắt văn bản, nhận dạng đối tượng từ hình ảnh, định vị hình ảnh trên bản đồ hoặc bằng cách kết hợp hình ảnh hai chiều để tạo mô hình ba chiều. Cuối cùng, AI có thể được sử dụng để phân tích hành vi của các đối tượng quan tâm nhằm rút ra những hiểu biết sâu sắc về hoạt động của họ, chẳng hạn như xác định chức năng của tòa nhà dựa trên phân tích mô hình cuộc sống và cũng có thể dự đoán các sự kiện và hoạt động trong tương lai.⁸

Tự động hóa các quy trình của tổ chức.

Các lực lượng vũ trang, giống như tất cả các tổ chức lớn, dựa vào một số lượng lớn các quy trình tổ chức, hành chính và quản lý dữ liệu để hoàn thành mục tiêu của họ. Đây là những hoạt động thường xuyên và thường lặp đi lặp lại, có thể tạo ra khối lượng công việc đáng kể. Việc sử dụng AI để tự động hóa các nhiệm vụ này có khả năng giúp tiết kiệm và cải thiện hiệu quả, giải phóng thời gian của nhân viên để giải quyết.

⁸ Alexander Babuta, Marion Oswald và Ardi Janjeva: 'Trí tuệ nhân tạo và An ninh Quốc gia Vương quốc Anh. Những cân nhắc về chính sách'. Viện Dịch vụ Tổng nhất Hoàng gia, ngày 27 tháng 4 năm 2020. P11. <https://securityaffairs.files.wordpress.com/2020/04/ai-and-national-security.pdf>



Ý tưởng của các nghệ sĩ về việc triển khai một đàn máy bay không người lái Gremlin. Tín dụng: DARPA

những vấn đề phức tạp hơn. Những hoạt động này bao gồm các yếu tố quản lý nhân sự, hậu cần, quản lý tài chính và kế toán.

Hoạt động mạng

An ninh mạng là lĩnh vực mà các hệ thống AI đang được sử dụng tích cực. Các mối đe dọa an ninh mạng như các cuộc tấn công bằng phần mềm độc hại phát triển nhanh chóng và đòi hỏi tốc độ phản ứng nhanh hơn nhiều so với tốc độ ra quyết định của con người. Hệ thống AI có thể chủ động xác định hoạt động đáng ngờ và ứng phó với các cuộc tấn công mạng trong thời gian thực. Bằng cách quét các mẫu hành vi đáng ngờ cũ ng như các mã độc hại tiềm ẩn, hệ thống AI có thể phát hiện các tín hiệu cho thấy mối đe dọa mạng mới.

Tuy nhiên, AI cũ ng có thể được sử dụng để tham gia vào các cuộc chiến tranh mạng mang tính tấn công. AI có thể được sử dụng để xác định các điểm yếu trong hệ thống phòng thủ mạng cũ ng như thiết kế phần mềm độc hại mới. Trong 'Thử thách lớn trên mạng' do DARPA tổ chức năm 2016, các thí sinh đã phát triển thuật toán AI có thể tự động xác định và vá các lỗ hổng trong phần mềm của chính họ đồng thời tấn công vào điểm yếu của các đội khác. Các thuật toán có thể tìm và sửa các lỗ hổng bảo mật chỉ trong vài giây, so với nhiều tháng sử dụng các phương pháp an ninh mạng thông thường. Phần mềm được phát triển như một phần của thử thách có khả năng đảm nhận đồng thời cả vai trò tấn công và phòng thủ, cung cấp những khả năng mới cho người dùng trong chiến tranh mạng trong đời thực.⁹

⁹ Kelley M. Saylor: 'Trí tuệ nhân tạo và An ninh Quốc gia'. Quốc hội Dịch vụ nghiên cứu. Ý tôi là. P10.

- 10 John Keller: 'Không quân yêu cầu ngành công nghiệp cho chiến tranh điện tử nhận thức (EW) trí tuệ nhân tạo (AI) cho máy bay phản lực F-15'. Điện tử Quân sự và Hàng không Vũ trụ, ngày 15 tháng 3 năm 2021. <https://www.quan.suaerospace.com/computers/bai-viet/14199230/dien-tu-chien-tranh-ew-nhan-thuc-nhan-tao-tri-tue-ai>
- 11 Sara Reese Hedberg: 'DART: Cách mạng hóa việc lập kế hoạch hậu cần'. Hệ thống thông minh IEEE, tháng 5/tháng 6 năm 2002. <https://www.gwezn.net/docs/ai/2002-hedberg.pdf>

Chiến tranh điện tử

Các hoạt động quân sự và hầu như tất cả các hệ thống vũ khí đều phụ thuộc vào phổ điện tử cho nhiều chức năng khác nhau. Chúng bao gồm việc sử dụng tần số vô tuyến, sóng vi ba, radar và thông tin vệ tinh. Trong chiến tranh, các chiến binh nhằm mục đích đảm bảo quyền truy cập không bị cản trở vào phổ điện tử cho các lực lượng đồng minh và ngăn chặn kẻ thù truy cập vào phổ điện tử. Chiến tranh điện tử là hành động sử dụng năng lượng điện tử để kiểm soát phổ điện tử, tấn công kẻ thù và chặn các cuộc tấn công của kẻ thù.

Những người ủng hộ lập luận rằng AI có vai trò trong tất cả các yếu tố của chiến tranh điện tử, với khả năng đối phó với các mối đe dọa mới và bất ngờ cũng như tốc độ hoạt động nhanh chóng của nó mang lại lợi thế so với các kỹ thuật thông thường.

Trong vòng hai năm tới, Không quân Hoa Kỳ đang hướng tới việc áp dụng thuật toán học máy và AI nhận thức vào hệ thống tác chiến điện tử trên máy bay chiến đấu F-15. Lực lượng Không quân muốn cho phép các hệ thống của máy bay phản ứng nhanh chóng với các mối đe dọa đang nổi lên và cung cấp khả năng học hỏi và lập trình lại nhanh chóng cho các hệ thống.¹⁰

DARPA cũng đang nghiên cứu các chương trình phát triển công nghệ tác chiến điện tử tự động.

Các biện pháp đối phó bằng radar thích ứng (ARC) của DARPA và Học tập hành vi cho chiến tranh điện tử thích ứng (BLADE) đều nhằm mục đích áp dụng AI để mô tả mối đe dọa chiến tranh điện tử 'đang diễn ra' và đưa ra một biện pháp đối phó với nó.

Chỉ huy và kiểm soát và hỗ trợ quyết định

Hệ thống chỉ huy và kiểm soát hỗ trợ người chỉ huy tác chiến trong việc chỉ đạo các nhiệm vụ và giám sát các lực lượng được giao nhiệm vụ, đồng thời giúp trình bày thông tin cho người chỉ huy ở định dạng dễ hiểu để hỗ trợ việc ra quyết định. Các hệ thống AI từ lâu đã được quân đội Hoa Kỳ sử dụng cho những mục đích như vậy. Trong Chiến tranh vùng Vịnh năm 1991, Bộ Tư lệnh Vận tải Hoa Kỳ đã sử dụng Công cụ Phân tích và Lập kế hoạch lại Động (DART), được phát triển bởi BBN Systems and Technologies và ISX Corporation theo chương trình DARPA, để lập kế hoạch và giải quyết các vấn đề trong quá trình vận chuyển vật tư quân sự từ các căn cứ ở Châu Âu đến Trung Đông.¹¹

Trong tương lai gần, có thể sử dụng AI để tập hợp dữ liệu từ nhiều loại cảm biến thành một bức tranh hoạt động chung cho người chỉ huy. Hình ảnh hoạt động chung là một bản đồ hiển thị các đối tượng quan tâm trên chiến trường, chẳng hạn như vị trí của lực lượng bạn và quân địch, cơ sở hạ tầng quan trọng và các thông tin hoạt động liên quan khác.

Lĩnh vực AI quân sự cũng đang nỗ lực phát triển các hệ thống có thể cung cấp hỗ trợ quyết định cho người chỉ huy. Các hệ thống hỗ trợ quyết định dựa trên AI đã được sử dụng trong các ứng dụng như chẩn đoán y tế để phân tích thông tin và đề xuất các hành động tiềm năng cho người vận hành.

Với trọng tâm là quân sự, những hệ thống như vậy sẽ nhằm mục đích giúp người chỉ huy ứng phó với các sự kiện đang diễn ra trên chiến trường, dựa trên phân tích dữ liệu theo thời gian thực và kiến thức về học thuyết chiến đấu của đối phương. Hệ thống có thể trình bày một menu các hành động có thể xảy ra cho người chỉ huy, kèm theo chỉ dẫn và giải thích về hậu quả có thể xảy ra của mỗi hành động. Mặc dù sự phán xét của con người sẽ là cần thiết

Trong việc ra quyết định chỉ huy và kiểm soát trong một thời gian, tốc độ và năng lực của các công cụ dựa trên AI có thể giúp loại bỏ các nhiệm vụ ngoại vi, cho phép người chỉ huy tập trung hoàn toàn vào các nhiệm vụ mà con người làm tốt hơn máy móc.¹²

Các lực lượng vũ trang của Vương quốc Anh rất muốn khám phá việc sử dụng các hệ thống như vậy để hỗ trợ các hoạt động của họ. Vào mùa hè năm 2021, Quân đội Anh tuyên bố đã sử dụng AI lần đầu tiên trong cuộc tập trận Bảo mùa xuân của NATO ở Estonia. Các binh sĩ tham gia cuộc tập trận đã sử dụng công cụ AI phân tích dữ liệu để cung cấp thông tin về môi trường và địa hình xung quanh, hỗ trợ lập kế hoạch cho các hoạt động quân sự và quản lý các quy trình chỉ huy và kiểm soát.¹³

Xe tự hành và robot quân sự

Máy bay và phương tiện điều khiển từ xa đã được sử dụng hiệu quả trong chiến đấu trong hai thập kỷ qua và AI hiện đang được sử dụng ngày càng nhiều để cho phép các hệ thống như vậy hoạt động tự chủ. Các ứng dụng AI trong lĩnh vực này tương tự như các ứng dụng đang được phát triển cho xe không người lái trong khu vực dân dụng, sử dụng cảm biến và phần mềm AI để nhận biết môi trường, nhận biết chướng ngại vật, kết hợp dữ liệu cảm biến, điều hướng và liên lạc với các phương tiện khác.

Công nghệ hiện đã đủ trưởng thành để các phương tiện tự hành quân sự sắp được các cường quốc quân sự lớn triển khai. Bộ Quốc phòng đã tài trợ cho Horiba Mira để

¹² Raúl Valencia: 'Trí tuệ nhân tạo trong Hệ thống Chỉ huy và Kiểm soát'. Blog GMV, ngày 16 tháng 6 năm 2020. <https://www.gmv.com/blog-gmv/lingu/en/artificial-intelligence-in-command-and-control-systems/>

¹³ 'Trí tuệ nhân tạo lần đầu tiên được sử dụng trong hoạt động của Quân đội.' Bộ của Quốc phòng, ngày 5 tháng 7 năm 2021. <https://www.gov.uk/Government/news/artificial-intelligence-used-on-army-operation-for-the-first-time>



Chú chó Hammer của Lực lượng Không quân Hoa Kỳ chạm trán với robot có chân bốn tự động Ghost Robotics Vision 60 tại Căn cứ Không quân Scott, tháng 12 năm 2020. Nguồn tin dụng: DoD

phát triển các phương tiện tự hành quân sự và hệ thống định vị để cung cấp vật tư cho các lực lượng tiền tuyến trong chiến đấu. Công ty đã được trao hợp đồng cung cấp ba phương tiện mặt đất không người lái tự hành Viking của mình để thử nghiệm với Quân đội Anh bắt đầu vào năm 2020.¹⁴

Máy bay không người lái có khả năng bay tự động cũng đang được phát triển. Phòng thí nghiệm nghiên cứu của Lực lượng Không quân Hoa Kỳ đã phát triển và thử nghiệm Hệ thống lái tự động Skyborg, một hệ thống phần mềm có thể thể hiện các khả năng hàng không cơ bản và đáp ứng các lệnh điều hướng, dự kiến có khả năng điều khiển máy bay không người lái trung thành. Không quân Hoa Kỳ được cho là có kế hoạch đưa máy bay chạy bằng động cơ Skyborg vào hoạt động vào năm 2023.¹⁵

Trong môi trường biển, Hải quân Hoàng gia đã tiến hành các cuộc thử nghiệm nâng cao với tàu quét mìn tự động và cũng đang thử nghiệm Madfox (Maritime Demonstrator For Operational experimentation), một tàu mặt nước không có người lái có thể thực hiện công việc giám sát và bảo vệ lực lượng theo cách tự động tự như khái niệm người chạy cánh trung thành. ¹⁶ Hải quân cũng đang thử nghiệm Phương tiện lặn không người lái cực lớn (XLUV).¹⁷

Bầy đàn

Phần mềm AI cũng đang được phát triển để cho phép các hệ thống tự động hoạt động như một bầy đàn thông minh được kết nối với nhau. Lấy cảm hứng từ đàn côn trùng, những cỗ máy tạo đàn có thể phối hợp hoạt động để áp đảo đối thủ. Các bầy đàn hoạt động tự chủ, không có sự kiểm soát trung tâm và các đơn vị thành phần riêng lẻ có thể cảm nhận được môi trường địa phương của chúng và các thành viên khác trong nhóm. Bầy đàn và hợp tác với các thành viên khác để thực hiện

một nhiệm vụ. MoD đã tài trợ cho một tập đoàn do Blue Bear Systems Research dẫn đầu, bao gồm Plextek DTS, IQHQ, Airbus và Đại học Durham, để phát triển công nghệ bầy đàn. Blue Bear đã phát triển một hệ thống chỉ huy và kiểm soát có khả năng quản lý một đàn và xử lý đồng thời các nhiệm vụ khác nhau. Trong các thử nghiệm, hệ thống có thể điều khiển một đàn gồm 20 máy bay không người lái, bao gồm 5 loại và kích cỡ khác nhau của máy bay không người lái cánh cố định với khả năng hoạt động khác nhau.¹⁸

Hệ thống vũ khí tự động gây chết người (LAWS)

LUẬT là hệ thống vũ khí được hỗ trợ bởi AI có khả năng xác định, tấn công và tiêu diệt mục tiêu mà không cần sự tương tác của con người. Chúng dựa trên sự kết hợp giữa hệ thống cảm biến giám sát môi trường xung quanh, hệ thống AI có thể xác định vật thể là mục tiêu tiềm năng và quyết định xem có nên tấn công nó hay không và vũ khí có thể tiêu diệt mục tiêu. Các công nghệ nền tảng mà khi kết hợp lại sẽ đóng vai trò là 'khối xây dựng' cho hệ thống vũ khí tự động sát thương đang phát triển với tốc độ chóng mặt.

Các hệ thống vũ khí đã tồn tại trong đó thuật toán đưa ra quyết định bắn chứ không phải con người. Những hệ thống có thể tiếp cận và tấn công mục tiêu một cách tự động hầu hết là các hệ thống phòng thủ, chẳng hạn như hệ thống phòng không hoặc robot canh gác SGR-A1 của Hàn Quốc.¹⁹ Các hệ thống này được thiết kế để vận hành dưới sự giám sát của con người và tự động khai hỏa trong các tình huống có thời gian giao tranh gấp. Được coi là quá ngắn để con người có thể đáp ứng. Tuy nhiên, các loại vũ khí tấn công có khả năng tiếp cận và tấn công mục tiêu một cách tự động cũng đang xuất hiện. Đạn bắn và quên, chẳng hạn như MBDA

14 Máy gia tốc quốc phòng và an ninh (DASA) Đánh giá thứ 09 năm 2019/2020, op cit.

15 Kyle Mizokami: 'Bộ não AI của Không quân vừa bay lần đầu tiên'. Cơ học phổ biến, ngày 13 tháng 5 năm 2021. <https://www.popularmechanics.com/military/hàng-không/a36412460/air-force-ai-brain-first-flight-skyborg-details/>

16 'Hải quân Hoàng gia lần đầu tiên có máy bay không người lái hệ thống quét mìn'. Bộ Quốc phòng, Thiết bị và Hỗ trợ Quốc phòng, ngày 5 tháng 5 năm 2018. <https://www.gov.uk/chính-phủ/tin-tức/royal-navy-gets-hệ-thống-quét-mìn-không-người-lái-đầu-tiên>

'Tàu tự hành mới được giao cho Hải quân Hoàng gia'. Hải quân Hoàng gia, ngày 26 tháng 3 năm 2021. <https://www.royalnavy.mod.uk/tin-tức-và-hoạt-động-mới-nhất/tin-tức/2021/tháng-ba/26/210326-madfox-tàu>

17 'Những cái tiến sẽ được thử nghiệm trên tàu ngầm tự hành tiên phong'. Máy gia tốc Quốc phòng và An ninh và Phòng thí nghiệm Khoa học và Công nghệ Quốc phòng, ngày 16 tháng 2 năm 2021. <https://www.gov.uk/chính-phủ/tin-tức/sự-đổi-mới-sẽ-được-thử-nghiệm-trên-tàu-ngầm-tiên-phong-tự-trị>

18 'Khái niệm máy bay không người lái bay lơ lửng gần với thực tế hơn'. Phòng thí nghiệm Khoa học và Công nghệ Quốc phòng và Máy gia tốc Quốc phòng và An ninh. Ngày 28 tháng 1 năm 2021. <https://www.gov.uk/Government/tin-tức/bầy-máy-bay-không-người-lái-khái-niệm-truồi-gần-với-thực-tế-hơn>

19 Ingvild Bode và Tom Watts: 'Kiểm soát con người ít ý nghĩa hơn: Bài học từ Hệ thống phòng không về kiểm soát con người có ý nghĩa đối với cuộc tranh luận về AWS', op cit.

Alexander Velez-Green: 'Tiểu luận về chính sách đối ngoại: Linh gác Hàn Quốc - Một "Robot sát thủ" để ngăn chặn chiến tranh'.

Lawfare, ngày 1 tháng 3 năm 2015. <https://www.lawfareblog.com/foreign-policy-essay-south-korean-sentry%E2%80%94killer-robot-prevent-war>

Tên lửa Brimstone đư ợc RAF sử dụng và đạn l ắng vàng IAI Harop có thể tự chọn mục tiêu và thể hiện nhiều tính năng của một vũ khí tự động sát thủ ợng.

Hoạt động thông tin

Các nền tảng truyền thông xã hội ngày càng đư ợc sử dụng làm nguồn tin tức và thông tin, tạo cơ hội cho những kẻ có mục đích xấu truyền bá thông tin sai lệch và gây hiểu lầm nhằm mục đích tạo ra sự chia rẽ và xung đột, thao túng các quy trình dân chủ và nhắm mục tiêu vào các cá nhân để cực đoan hóa hoặc khuyến khích họ. không tuân theo h ư ớng dẫn. Các hoạt động này có thể đư ợc tiến hành song song với các hoạt động quân sự 'trên bộ' để giúp đạt đư ợc các mục tiêu quân sự.

Các bot internet hỗ trợ AI - các ứng dụng phần mềm chạy các tác vụ tự động đơn giản và lặp đi lặp lại - có thể tạo điều kiện thuận lợi cho các chiến dịch như vậy bằng cách tạo danh tính trực tuyến giả mạo và truyền bá thông tin nhanh hơn nhiều so với khả năng của một ngư ời. Các ứng dụng AI cũ ng có thể đư ợc sử dụng để khai thác dữ liệu từ mạng xã hội nhằm tạo ra "mô hình cuộc sống" kỹ thuật số cho các cá nhân - bao gồm các quan chức chính phủ, chính trị gia và thành viên của lực l ư ợng vũ trang - nhằm mục đích c ư ỡng chế. Hình ảnh do AI tạo ra có thể tạo ra ph ư ợng tiện tổng hợp - cảnh quay hình ảnh và video ngày càng chân thực, hay còn gọi là 'giả mạo sâu sắc' - mà cuối cùng có thể đư ợc sử dụng để hỗ trợ các hành vi xúc phạm trực tuyến nhằm mục đích đánh l ử a và thao túng.



AI đư ợc các c ư ờng quốc quân sự thế giới coi là một cách để cách mạng hóa chiến tranh và giành lợi thế trư ớc kẻ thù.

Các ứng dụng quân sự

của AI đã đư ợc sử dụng hàng ngày trong vài thập kỷ qua và các hệ thống mới với những đặc điểm đáng lo

ngại đang nhanh chóng đư ợc triển khai



Phát triển AI trong quân đội Anh

20 'Chiến lược AI quốc gia'. Văn phòng cho Trí tuệ nhân tạo, Phòng Kỹ thuật số, Văn hóa, Truyền thông & Thể thao và Phòng Chiến lược Kinh doanh, Năng lực & Công nghiệp. Ngày 22 tháng 9 năm 2021. <https://www.gov.uk/government/publications/national-ai-strategy>

21 'Nước Anh toàn cầu trong thời đại cạnh tranh: Đánh giá tổng hợp về chính sách an ninh, quốc phòng, phát triển và đối ngoại'. Văn phòng Nội các, ngày 16 tháng 3 năm 2021. P7. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/975077/Global_Britain_in_a_Competitive_Age_the_Integrated_Review_of_Security__Quốc_phòng_Phát_triển_và_Đối_ngoại_Chính_sách.pdf

22 Bộ Quốc phòng: 'Phòng thủ trong một Thời đại cạnh tranh'. Văn bản chỉ đạo CP 411, tháng 3 năm 2021. https://assets.publishing.service.gov.uk/government/uploads/attachment_data/file/974661/CP411_Defence_Command_Plan.pdf

23 'Quốc phòng và an ninh'. Viện Alan Turing. <https://www.turing.ac.uk/research/ai-for-defence>

24 'Chiến lược Khoa học và Công nghệ 2020'. Bộ Tư lệnh Quốc phòng. Ngày 19 tháng 10 năm 2020. P15. https://www.gov.uk/government/uploads/attachment_data/file/974661/CP411_Defence_Command_Plan.pdf

25 'Bài phát biểu tại hội nghị của Tư lệnh Bộ Tư lệnh Chiến lược RUSI'. Bộ Quốc phòng, Bộ Tư lệnh Chiến lược và Tư lệnh Sir Patrick Sanders KCB CBE DSO ADC Gen, ngày 26 tháng 5 năm 2021. <https://www.gov.uk/government/speeches/sir-patrick-sanders-rusi-ai-strategy>

Chính phủ Anh rõ ràng rất coi trọng việc phát triển Trí tuệ nhân tạo.

Vào tháng 9 năm 2021, chính phủ Vương quốc Anh đã công bố Chiến lược AI quốc gia, trong đó đặt ra kế hoạch 10 năm "đưa nước Anh trở thành Siêu cường AI toàn cầu". Chiến lược này đưa ra các đề xuất đầu tư vào lĩnh vực AI, đặt AI trở thành xu hướng chủ đạo của nền kinh tế Vương quốc Anh và giới thiệu nó trên tất cả các lĩnh vực kinh tế và khu vực của Vương quốc Anh.

Cách tiếp cận của Vương quốc Anh đối với quy định về AI sẽ được trình bày trong sách trắng sẽ xuất bản vào đầu năm 2022.20

Đánh giá Tích hợp năm 2021 đặt ra mong muốn của chính phủ là "được công nhận là Siêu cường Khoa học và Công nghệ. và thiết lập được lợi thế dẫn đầu trong các lĩnh vực quan trọng như trí tuệ nhân tạo".21 Tài liệu Chỉ huy Quốc phòng đi kèm với Đánh giá Tích hợp nêu rõ rằng MOD dự định đầu tư 6,6 tỷ bảng Anh trong bốn năm tới vào nghiên cứu và phát triển quốc phòng, tập trung vào các công nghệ mới nổi trong trí tuệ nhân tạo, hệ thống tự động hỗ trợ AI, hệ thống mạng, không gian và năng lực định hướng.22 Một số nhóm đang thúc đẩy sự phát triển của quân đội Các hệ thống AI trong MoD, bao gồm các đơn vị trong mỗi lực lượng vũ trang, Phòng thí nghiệm Khoa học và Công nghệ Quốc phòng (DSTL) và Thiết bị và Hỗ trợ Quốc phòng (DES).

Quốc phòng và an ninh cũng là lĩnh vực nghiên cứu cốt lõi của Viện Alan Turing, viện quốc gia về khoa học dữ liệu và trí tuệ nhân tạo của Vương quốc Anh.23

Bộ Tư lệnh Chiến lược (chịu trách nhiệm về hệ thống máy tính và tình báo quốc phòng), DSTL và GCHQ được liệt kê là đối tác của chương trình quốc phòng và an ninh của Viện, được tài trợ bởi MoD và GCHQ.

Bộ Quốc phòng đã thành lập Đơn vị Tự chủ và Trí tuệ Nhân tạo Quốc phòng để giải quyết các vấn đề chính sách, đồng thời lên kế hoạch xuất bản Chiến lược Tự chủ và Trí tuệ Nhân tạo vào mùa hè năm 2021 để cho phép áp dụng nhanh chóng các công nghệ.24 Định hướng của chiến lược đã được Đại tư lệnh vạch ra Patrick Sanders, Tư lệnh Bộ Tư lệnh Chiến lược, trong bài phát biểu tại Viện Dịch vụ Thống nhất Hoàng gia vào tháng 5 năm 2021:

"Nó sẽ bắt đầu bằng việc tích hợp các công nghệ kỹ thuật số hiện có - ví dụ như sử dụng máy học và tự động hóa để hỗ trợ phân tích Thông minh. Nó sẽ được kích hoạt bằng cách cải thiện cơ sở hạ tầng kỹ thuật số của chúng tôi - xử lý sống kỹ thuật số - với chiến lược dữ liệu cho phép quản lý dữ liệu, chia sẻ và khai thác dữ liệu, các dịch vụ đám mây tại Secret và Above Secret cũng như kiến trúc mạng chung.

Nó sẽ dẫn đến việc đầu tư vào KH&CN nhiều hơn trong quan hệ đối tác với DSTL và thử nghiệm để đảm bảo sự phát triển có trách nhiệm của các hệ thống tự động và hỗ trợ AI."25

- 26 'Chuyên môn do DE&S lãnh đạo sẽ cách mạng hóa phát triển robot quân sự của Vũ ơng quốc Anh'. Bộ Quốc phòng và Hỗ trợ Trang thiết bị Quốc phòng'21 tháng 5 năm 2021. <https://www.gov.uk/ chinh-phu/tin-tuc/ides-led-expertise-set-to-cach-mang-phat-trien-of-uk-quan-doi-ngu-oi-may>
- 27 Richard Brantingham: 'NELSON Tiêu chuẩn - Một hệ thống thiết kế cho Hải quân Hoàng gia'. Bộ Quốc phòng, ngày 18 tháng 12 năm 2019. <https://defencedigital.blog.gov.uk/2019/12/18/nelson-standards-creating-royal-navy-apps-with-a-consistent-look-and-feel/>
- 'Hải quânX'. Hải quân Hoàng gia. <https://www.royalnavy.mod.uk/navyx>
- 28 'Lực lượng thử nghiệm chiến đấu được tiết lộ'. Bộ Quốc phòng, ngày 23 tháng 3 năm 2021. <https://www.gov.uk/ Government/news/ chiến-đấu-thử-nghiệm-lực-lượng-ra-mắt>
- 29 'Khoa học cho phép binh lính và phi hành đoàn hợp tác với máy móc thông minh ở Salisbury Plain'. Phòng thí nghiệm Khoa học và Công nghệ Quốc phòng và Máy gia tốc Quốc phòng và An ninh, ngày 23 tháng 9 năm 2020. <https://www.gov.uk/ Government /news/science-enables-soldiers-and-aircrew-to-partner-with-smart-machines-on-salisbury-plain>
- Hoàn thiện 30 '30 triệu bảng Anh cho chiếc tàu chử a được lắp ráp đầu tiên của Vũ ơng quốc Anh máy bay chiến đấu'. Bộ Quốc phòng, ngày 25 tháng 1 năm 2021. <https://www.gov.uk/ Government/ tin-tuc/30m-tiem-cho-uks-dau-tien-khong-co-ngu-oi-lai-may-bay-chien-dau>
- 31 'Phòng thí nghiệm AI hàng đầu được công bố là Phòng thí nghiệm Quốc phòng Bộ tư ơng chủ trì cuộc gặp đầu tiên giữa các nhà đổi mới quốc phòng Anh và Mỹ' Bộ Quốc phòng, ngày 22 tháng 5 năm 2018. <https://www.gov.uk/ chinh-phu/tin-tuc/flagship-ai-lab-announced-as-defence-secretary-hosts-first-meet-between-british-and-american-defence-innovators>
- 32' Khái niệm chung Ghi chú 17/2. Tư ơng lai của Chỉ huy và Kiểm soát'. Bộ Quốc phòng, ngày 8 tháng 9 năm 2017. https://assets.publishing.service.gov.uk/ Government/uploads/system/tai-len/attachment_data/file/643245/concept_uk_future_c2_jcn_2_17.pdf
- 'Lư u ý chung về khái niệm 18/1. Hợp tác giữa ngư ời và máy'. Bộ Quốc phòng, ngày 18 tháng 5 năm 2018. https://assets.publishing.service.gov.uk/ chinh-phu/tai-len/he-thong/tai-len/attachment_data/file/709359/20180517-khai-niem_uk_human_machine-teaming_jcn_1_18.pdf

Tổ chức mua sắm thiết bị của Bộ Quốc phòng, Thiết bị và Hỗ trợ Quốc phòng (DES), đã thành lập một đơn vị đặc biệt để phát triển công nghệ tự hành.

Trung tâm Chuyên môn về Robot Viễn chinh đã đư ợc DES thành lập với sự cộng tác của Quân đội Anh để tập hợp các chuyên gia về robot và hệ thống tự động từ khắp chính phủ, học viện và ngành công nghiệp nhằm đánh giá "các công nghệ chử a đư ợc khám phá, có rủi ro cao như ng đang phát triển nhanh chóng".26

Mỗi lực lượng vũ trang đều có đơn vị thử nghiệm riêng để phát triển các công nghệ quân sự mới, chẳng hạn như AI. Hải quân Hoàng gia đã thành lập cơ quan phát triển phần mềm của riêng mình, NELSON, và đã thành lập NavyX, một 'Máy gia tốc tự động và sát thương' nhằm phát triển tàu mặt nư ớc tự hành và máy bay không ngư ời lái.27

Quân đội đang giới thiệu Lực lượng Thử nghiệm Chiến tranh, dựa trên một tiểu đoàn bộ binh với các thành phần đư ợc rút ra từ khắp Quân đội, lực lượng này sẽ thử nghiệm công nghệ mới và sự tích hợp của nó.28 Trong mư ời năm qua, Quân đội đã tiến hành các sự kiện Thử nghiệm Chiến tranh tư ơng xuyên của Quân đội trên Đồng bằng Salisbury để thử nghiệm các công nghệ mới đư ợc phát triển bởi DSTL và ngành công nghiệp. Chúng bao gồm "các hệ thống không ngư ời lái bán tự động" và các hệ thống giúp giảm tải nhận thức cho ngư ời vận hành hệ thống "trong khi không tự chủ hoàn toàn".29

Lực lượng Không quân Hoàng gia có Văn phòng Năng lực Nhanh, kể từ năm 2017, cơ quan này đã khám phá các lựa chọn cho Máy bay chiến đấu hạng nhẹ mới, giá cả phải chăng (LANCA), một loại máy bay chiến đấu không ngư ời lái chi phí thấp đư ợc thiết kế để bay ở tốc độ cao cùng với máy bay có ngư ời lái với tư cách là 'ngư ời chạy cánh trung thành'. ' để cung cấp hỗ trợ, bảo vệ và thực hiện các nhiệm vụ độc lập như trinh sát, tác chiến điện tử hoặc nhiệm vụ chiến đấu. Spirit Aerosystems đã đư ợc trao hợp đồng trị giá 30 triệu bảng Anh để thực hiện 'Dự án Mosquito', nhằm mục đích

đưa ra chử ơng trình bay thử nghiệm vào cuối năm 2023 với hy vọng máy bay không ngư ời lái sẽ đư ợc đưa vào sử dụng vào cuối thập kỷ này.30

Phòng thí nghiệm Khoa học và Công nghệ Quốc phòng (DSTL), bộ phận nghiên cứu của Bộ Quốc phòng, đã thành lập Phòng thí nghiệm AI, dự định trở thành lá cờ đầu cho AI, học máy và khoa học dữ liệu, vào tháng 5 năm 2018 tại trụ sở Porton Down. Phòng thí nghiệm AI nhằm mục đích tăng cường và đẩy nhanh việc ứng dụng các công nghệ liên quan đến AI vào các thách thức quốc phòng và an ninh.31

DSTL cũ ng hỗ trợ Công cụ tăng tốc quốc phòng và an ninh (DASA), một nguồn tài trợ cho sự đổi mới trong quốc phòng và an ninh.

Cho đến nay, điều còn thiếu trong tất cả những điều này là việc xem xét các vấn đề đạo đức liên quan đến việc sử dụng AI. Đánh giá tích hợp và các tuyên bố khác của chính phủ không còn nghi ngờ gì nữa rằng chính phủ rất coi trọng các ứng dụng quân sự của AI và có ý định chạy đua với sự phát triển của nó. Tuy nhiên, mặc dù các ấn phẩm phác thảo học thuyết về việc sử dụng các hệ thống tự động đã đư ợc xuất bản,32 cho đến nay Bộ Quốc phòng vẫn giữ im lặng về khuôn khổ đạo đức quản lý việc sử dụng AI và các hệ thống tự động của mình, mặc dù đã đưa ra các quyết định quan trọng về việc sử dụng quân sự trong tư ơng lai. AI.

Rủi ro do hệ thống AI quân sự gây ra

Mục tiêu sử dụng AI và hệ thống tự hành cho mục đích quân sự đang gây tranh cãi. Ngay cả các nguồn tin quân sự cũng thừa nhận rằng có những rủi ro đáng kể và sẽ là không khôn ngoan nếu lao vào hành động bừa bãi.

phát triển và triển khai các khả năng này.

Rõ ràng có các yếu tố rủi ro khác nhau liên quan đến từng ứng dụng quân sự khác nhau của AI được mô tả trong phần trước. Thuật toán phân loại dữ liệu như một phần của hoạt động hỗ trợ tại trụ sở Bộ Quốc phòng sẽ đặt ra các cấp độ vấn đề và mối lo ngại khác, đồng thời yêu cầu mức độ giám sát khác với hệ thống vũ khí tự động.

Nhìn về phía trước trong khoảng thời gian khoảng mười năm tới, các mối đe dọa do triển vọng cực đoan hơn như xa vời hơn về việc phát triển hệ thống AI siêu thông minh gây ra rủi ro hiện hữu cho nhân loại có thể được giảm bớt.

Tuy nhiên, các hệ thống AI hiện đang được phát triển chắc chắn gây ra những mối đe dọa đối với tính mạng, nhân quyền và hạnh phúc. Khi cạnh tranh để đạt được lợi ích quân sự của AI trong thời gian sớm nhất, các quốc gia có thể không áp dụng các biện pháp phòng ngừa thích hợp, dẫn đến một 'cuộc đua xuống đáy'. Do đó, các ứng dụng quân sự của AI có thể làm giảm, thay vì tăng cường tính bảo mật.

Những rủi ro do hệ thống AI quân sự gây ra có thể được nhóm thành ba loại: đạo đức và pháp lý, hoạt động và chiến lược.³³

³³ Forrest E. Morgan, Benjamin Boudreaux, Andrew J. Lohn, Mark Ashby, Christian Curriden, Kelly Klima, Derek Grossman: 'Ứng dụng quân sự của trí tuệ nhân tạo: Mối quan tâm về đạo đức trong một thế giới không chắc chắn', op cit. P30.

Rủi ro đạo đức và pháp lý

Tuân thủ luật chiến tranh

Hệ thống AI và robot quân sự phải có khả năng được sử dụng phù hợp với luật nhân đạo quốc tế và luật nhân quyền quốc tế, vốn chi phối việc tiến hành chiến tranh. Ở cấp độ thực tế, chưa rõ liệu các hệ thống robot, đặc biệt là vũ khí tự động, có thể đáp ứng các tiêu chuẩn do luật chiến tranh đặt ra về việc đưa ra các quyết định gây chết người và bảo vệ những người không tham chiến hay không.

Trong chiến tranh, người chỉ huy lập kế hoạch và quyết định tấn công có nghĩa vụ pháp lý để phân biệt người tham chiến và người không tham chiến; đảm bảo thiệt hại gây ra tương ứng với mục tiêu quân sự và có biện pháp phòng ngừa để bảo vệ dân thường. Tổn thất nhân mạng của dân thường do 'thiệt hại tài sản thể chấp' chỉ hợp pháp nếu nó tương xứng với mục tiêu quân sự và các chỉ huy có thể phải chịu trách nhiệm về bất kỳ hành vi vi phạm nào.

Tuy nhiên, những phán đoán cần thiết để phân biệt người tham chiến với người không tham chiến, đánh giá mức độ tương ứng của một cuộc tấn công và đưa ra biện pháp phòng ngừa, phần lớn là những quyết định định tính dựa trên kinh nghiệm, ý thức chung và sự hiểu biết về bối cảnh mà hệ thống robot không có. Kỹ thuật và



Các tòa nhà bị đánh bom trong các cuộc tấn công vào Raqqa, Syria vào năm 2018. Việc sử dụng AI và hệ thống tự động để tiến hành chiến tranh có thể khiến không thể quy trách nhiệm về thứ ông vong cho dân thường và tội ác chiến tranh. Nguồn: Tổ chức Ân xá Quốc tế

những thách thức về lập trình trong việc thiết kế các hệ thống máy tính có thể hoạt động tuân thủ luật chiến tranh trong các môi trường phức tạp, lộn xộn và năng động vẫn còn rất lớn. Vì những lý do này, có vẻ như sự kiểm soát của con người trong việc lựa chọn và tấn công các mục tiêu sẽ tiếp tục cần thiết để đảm bảo tuân thủ luật xung đột vũ trang.

Các hệ thống vũ khí tự động cũng có thể không vượt qua được bài kiểm tra về việc tuân thủ “các nguyên tắc nhân đạo và yêu cầu của luật tâm công chúng” - ‘Điều khoản Martens’ có trong Công ước La Hay năm 1899. Vì máy móc không thể thể hiện lòng trắc ẩn hoặc sự đồng cảm, nên người ta lập luận rằng chúng thiếu sự kiềm chế vốn có của con người đối với việc giết chóc và không thể vượt qua thách thức xem xét các khía cạnh đạo đức của việc giết chóc.

Trách nhiệm giải trình

Đã có những lo ngại rằng việc sử dụng hệ thống AI có thể dẫn đến “khoảng cách về trách nhiệm giải trình” nếu xảy ra vi phạm luật nhân đạo. Không rõ ai sẽ bị giữ

chịu trách nhiệm nếu mọi việc diễn ra không như ý muốn - tuy nhiên sẽ vô nghĩa khi trừng phạt một máy tính nếu nó hoạt động không thể đoán trước và kết quả là phạm tội ác chiến tranh. Sự phức tạp của hệ thống AI có thể gây khó khăn cho việc quyết định liệu tội phạm có phải là trách nhiệm của người vận hành, lập trình viên hay nhà sản xuất hay chỉ là một tai nạn khó lường mà cuối cùng không ai có lỗi.

Nhân quyền và quyền riêng tư

Hệ thống AI cũng đặt ra những mối đe dọa tiềm tàng đối với nhân quyền và quyền riêng tư cá nhân. Các tính năng của AI cho phép khai thác dữ liệu, giám sát liên tục, phân tích các đặc điểm vật lý như nhận dạng khuôn mặt hoặc dáng đi và xâm nhập vào hoạt động trực tuyến

cung cấp cho các chính phủ và doanh nghiệp lớn phạm vi rộng để giám sát và kiểm soát dân số cũ ng như lạm dụng nhân quyền.

Các hoạt động thông tin nhằm truyền bá thông tin sai lệch và tìm cách thao túng dư luận hoặc tạo chia rẽ cũ ng có thể gây ra những tác hại lớn hơn như hạ thấp tầm quan trọng của các sự kiện và thông tin khách quan trong quá trình ra quyết định và làm xói mòn niềm tin cần thiết cho hoạt động dân chủ. Sự phát triển của công nghệ AI cho mục đích quân sự rõ ràng sẽ làm trầm trọng thêm rủi ro về những tác động không mong muốn này.

Rủi ro hoạt động

Rủi ro hoạt động là những rủi ro liên quan đến sự thất bại của AI trong các ứng dụng quân sự theo những cách không lường trước được. Một số rủi ro này có liên quan đến các dạng sai lệch khác nhau.³⁴

Các nguồn sai lệch kỹ thuật

Các hệ thống tự trị phải được 'huấn luyện' bằng cách sử dụng một bộ dữ liệu để so sánh dữ liệu từ thế giới thực mà nó hoạt động. Hệ thống AI chỉ hoạt động tốt khi dữ liệu đào tạo của chúng hoạt động tốt và một lượng nhỏ dữ liệu đào tạo bị hỏng có thể ảnh hưởng lớn đến hiệu suất của hệ thống. Ngay cả trong điều kiện tốt nhất, các hệ thống AI được đào tạo kém vẫn có thể tạo ra kết quả phân biệt đối xử với các nhóm cụ thể.³⁵

Một nguồn sai lệch khác được gọi là sai lệch xử lý và xảy ra khi một thuật toán tự bị sai lệch trong cách nó biến đổi dữ liệu. Những thành kiến cũ ng có thể phát sinh từ việc sử dụng hoặc triển khai hệ thống không phù hợp. Các hệ thống thuật toán được phát triển và huấn luyện cho các mục đích cụ thể trong

bối cảnh cụ thể và chắc chắn có một phạm vi hoạt động hẹp. Đưa hệ thống ra ngoài bối cảnh này và vượt quá khả năng đã được thử nghiệm của nó có thể dẫn đến hành vi không mong muốn và có thể nguy hiểm. Trong môi trường chiến đấu chắc chắn phức tạp và thay đổi nhanh chóng, các hệ thống tự hành sẽ phải đối mặt với những tình huống mới cản trở việc thu thập dữ liệu và không phù hợp với yêu cầu của chúng.

hoàn cảnh mà hệ thống được phát triển và thử nghiệm. Kết quả là họ sẽ phải chịu những thất bại không thể lường trước được.

Nguồn gốc của sự thiên vị con người

Sự thiên vị có thể xảy ra khi con người sử dụng sai hệ thống hoặc hiểu sai kết quả đầu ra của nó. Những thành kiến này liên quan đến mức độ tin cậy của người vận hành đối với độ tin cậy của hệ thống. Sự thiên vị có thể xảy ra khi người dùng không diễn giải chính xác đầu ra của hệ thống AI và do đó hành động sai. Nếu người dùng không hiểu chính xác ý nghĩa của đầu ra hoặc bị định hướng sai bởi các tính năng không đáng tin cậy của đầu ra thì sẽ mắc lỗi.

Một thách thức như vậy là sự tự tin quá mức vào đầu ra của một hệ thống tự trị. Trong Chiến tranh Iraq năm 2003, hệ thống phòng không Patriot tự động hóa cao đã bắn hạ một máy bay Tornado của Không quân Hoàng gia, giết chết phi hành đoàn khi hệ thống xác định không chính xác máy bay là kẻ thù và người điều khiển con người đã chấp nhận nhận dạng không chính xác và cho phép tham gia. Loại 'thiên vị tự động hóa' này xảy ra nếu người vận hành đặt quá nhiều niềm tin vào khả năng của hệ thống tự động.

Mặt khác, tai nạn cũ ng có thể phát sinh trong tình huống ngược lại, khi người vận hành không tin tưởng vào hệ thống và không đủ tin cậy vào quy trình tự động. Điều này rất có thể xảy ra với một hệ thống

³⁴ Xu hướng thuật toán và Vũ khí hóa các công nghệ ngày càng tự chủ'. Viện Nghiên cứu Giải trừ Quân bị của Liên Hợp Quốc, ngày 22 tháng 8 năm 2018. [https://unidir.org/files/publications/pdfs/thuật toán-thiên vị-và-vũ khí-ngày càng -tự trị-công nghệ-en-720.pdf](https://unidir.org/files/publications/pdfs/thuật%20toán-thiên%20vị-và-vũ%20khí-ngày%20càng-tự%20trị-công%20nghệ-en-720.pdf)

³⁵ Arthur Holland Michel: 'Đã biết Không biết'. Viện Nghiên cứu Giải trừ Quân bị của Liên Hợp Quốc, ngày 17 tháng 5 năm 2021. <https://unidir.org/known-unknowns>

đư ợc biết là tạo ra lỗi 'đư ợng tính giả'. Trong những tình huống như vậy, ngư ời vận hành có thể bỏ qua thông tin quan trọng do hệ thống cung cấp hoặc ghi đè các hành động của nó mà không có lý do chính đáng. Một ví dụ bi thảm về điều này là vụ bắn hạ máy bay chở khách Airbus A300, Chuyến bay 655 của Iran Air và khiến 290 ngư ời thiệt mạng vào ngày 3 tháng 7 năm 1988 bởi hệ thống phòng không Aegis tự động trên tàu chiến USS Vincennes. Trái ngư ợc với thông tin đư ợc hệ thống cung cấp, thủy thủ đoàn của tàu quyết định rằng tín hiệu radar đến từ một máy bay chiến đấu đang tấn công và quyết định bắn hạ nó.³⁶

Rủi ro về những sự cố như vậy càng gia tăng trong các hệ thống phức tạp đến mức đầu ra của chúng không thể giải thích đư ợc và trong đó ngư ời vận hành không thể dễ dàng xác định lý do tại sao hệ thống của họ đư a ra những quyết định cụ thể hoặc đang hoạt động theo những cách cụ thể. Đồng thời, sự tự tin quá mức - tin rằng chúng ta hiểu rõ hơn về cách thức hoạt động của hệ thống so với thực tế - cũ ng có thể dẫn đến sai lầm trong việc hiểu lý do hành vi của hệ thống AI và dự đoán cách nó có thể hoạt động trong tương lai.

Thao túng với mục đích xấu

Các hệ thống AI quân sự, giống như tất cả các hệ thống nổi mạng, dễ bị tấn công từ các tác nhân độc hại, những kẻ có thể cố gắng gây nhiễu, hack hoặc giả mạo hệ thống. Các hệ thống không chỉ dễ bị tin tặc tấn công, những ngư ời có thể cố gắng truy cập trực tiếp để thao túng hoặc chiếm quyền kiểm soát hệ thống mà còn dễ bị tấn công bởi các kiểu tấn công khác. Chúng bao gồm các cuộc tấn công đầu độc dữ liệu, trong đó dữ liệu huấn luyện bị thao túng nhằm làm đảo lộn chức năng của hệ thống. Cũ ng có thể giả mạo một hệ thống bằng cách lừa nó mắc lỗi bằng cách đư a cho nó những dữ liệu giả mạo tại hiện trư ờng.

Tính năng hệ thống vốn có

Hệ thống máy học có thể là hệ thống 'hộp đen' mà con ngư ời không thể quan sát hoặc hiểu cách đư a ra quyết định. Không thể lập trình lại một hệ thống như vậy để thay đổi phư ơng thức hoạt động của nó. Nếu một hệ thống máy học đư ợc phép tiếp tục học theo cách không giám sát thì nó có khả năng phát triển, nghĩa là hệ thống đang hoạt động sẽ không giống với hệ thống đã đư ợc thử nghiệm và có thể đư a ra các quyết định khác.

Rủi ro chiến lược

Ngư ỡng

Mặc dù những ngư ời ủng hộ ứng dụng quân sự của AI cho rằng chúng có thể giảm nguy cơ gây tổn hại cho quân nhân, như ng điều này gây ra nguy cơ mới là các nhà lãnh đạo sẽ sử dụng hệ thống quân sự tự trị trong xung đột thay vì theo đuổi các lựa chọn phi quân sự để giải quyết khác biệt, hạ thấp ngư ỡng cho việc sử dụng vũ lực. Mặc dù có thể có ít binh sĩ hơn ở vùng chiến sự như ng dân cư vẫn sẽ hiện diện và sẽ tiếp tục đối mặt với rủi ro chiến tranh, chuyển chi phí chiến tranh từ binh lính sang dân trư ờng.

Quản lý leo thang

Việc sử dụng các hệ thống không phải của con ngư ời trong chiến tranh làm tăng khả năng leo thang thông qua một số cơ chế. Thứ nhất, nếu chiến tranh dễ xảy ra hơn do sử dụng AI và công nghệ tự hành, thì bản thân tần suất hành động quân sự ngày càng tăng sẽ làm tăng nguy cơ leo thang. Thứ hai, nếu các chỉ huy tin rằng việc sử dụng công nghệ tự động có nghĩa là binh lính của họ

36 Vincent Boulanin và Maaike Verbruggen: 'Lập bản đồ sự phát triển quyền tự chủ trong các hệ thống vũ khí'. Viện Nghiên cứu Hòa bình Quốc tế Stockholm, tháng 11 năm 2017. P40. https://www.sipri.org/sites/default/files/2017-11/siprireport_mapping_the_development_of_autonomy_in_wisystems_1117_1.pdf Elisa B. Kania: 'Con ngư ời phê phán yếu tố trong thời đại máy móc của chiến tranh'. Bản tin của các nhà khoa học nguyên tử, ngày 15 tháng 11 năm 2017. <https://thebulletin.org/quan-trong-con-nguoi-yếu-tổ-máy-age-warfare11277#>

sẽ đối mặt với ít tổn hại hơn trong xung đột, họ có thể bị cám dỗ chấp nhận rủi ro lớn hơn và hành động quyết liệt hơn, thúc đẩy sự leo thang. Cuối cùng, tốc độ thực hiện hành động quân sự - ở tốc độ máy móc, thay vì trong khung thời gian của con người - tạo ra nguy cơ xảy ra một 'tia sáng' ngoài ý muốn.

chiến tranh' có thể đột nhiên phát triển. Không gian cho việc cân nhắc và đàm phán sẽ giảm đi, dẫn đến khả năng leo thang nhanh chóng, đặc biệt là trong các tình huống khủng hoảng khi nguy cơ đánh giá sai hoặc hành động sai lầm có thể gây ra hậu quả nghiêm trọng.

Sự kết hợp của các công nghệ quân sự mới nổi có thể làm tăng nguy cơ leo thang hơn nữa. Công nghệ siêu thanh cho phép tên lửa di chuyển với tốc độ gấp nhiều lần tốc độ âm thanh, khi kết hợp với công nghệ ra quyết định AI, sẽ đặc biệt gây mất ổn định và nguy hiểm, đồng thời có thể giảm thời gian phản ứng đến mức ngăn chặn khả năng có bất kỳ sự can thiệp nào của con người.

Chiến tranh mạng có thể chứng kiến những bước phát triển tự động, trong đó các hệ thống phần mềm AI không chỉ có thể cung cấp khả năng phòng vệ trước một cuộc tấn công mạng mà còn khởi động một cuộc tấn công trả đũa tự động.³⁷

Chạy đua vũ trang và phổ biến vũ khí

Việc theo đuổi AI quân sự với mục đích giành lợi thế trước các đối thủ tiềm năng dường như đã gây ra cuộc chạy đua vũ trang, với các cường quốc và khu vực chạy đua để phát triển khả năng của mình nhằm đi trước các đối thủ. Cuộc cạnh tranh này không có mục tiêu cuối cùng tuyệt đối, chỉ có mục tiêu tương đối là vượt lên trên các đối thủ khác, nhưng nó làm tăng nguy cơ hiểu lầm và leo thang.

Nhiều ứng dụng AI có mục đích sử dụng kép, nghĩa là các công nghệ dân sự có thể được áp dụng cho mục đích quân sự và có thể

có sẵn từ các nguồn thương mại hoặc nguồn mở.

Các công nghệ khác nhau có thể được kết hợp để sử dụng với mục đích xấu. Khi năng lực được cải thiện và chi phí giảm xuống, AI và các công nghệ lưu trữ dữ liệu sẽ trở nên sẵn có hơn đối với các cường quốc nhỏ, các chủ thể phi nhà nước và các nhóm tội phạm.

Sự ổn định chiến lược

Nếu các hệ thống AI tiên tiến phát triển đến mức có thể dự đoán chiến thuật của kẻ thù hoặc việc triển khai lực lượng, điều này có thể gây ra hậu quả mất ổn định cao, đặc biệt là trong lĩnh vực răn đe hạt nhân. Nếu AI có thể dự đoán vị trí lực lượng của đối thủ, kẻ xâm lược có thể bị cám dỗ tiến hành một cuộc tấn công phủ đầu với ý định tiêu diệt chúng mà không sợ bị trả thù.

Ngay cả nhận thức rằng các lực lượng có thể dễ bị tổn thương theo cách này cũng có thể khuyến khích một quốc gia thực hiện đòn tấn công đầu tiên “sử dụng hoặc mất chúng” để tránh khả năng họ không thể sử dụng lực lượng của mình sau này trong một cuộc xung đột.

Kết luận: Tai nạn sẽ xảy ra

Tai nạn là không thể tránh khỏi trong các hệ thống phức tạp và điều này bao gồm cả các hệ thống tự động cơ bản. Khi các hệ thống trở nên phức tạp hơn và “liên kết chặt chẽ” với các hệ thống khác, tạo ra “hệ thống của các hệ thống”, nguy cơ tai nạn ngoài ý muốn sẽ tăng lên. Đối với các hệ thống tự động và AI phức tạp hoạt động ở tốc độ máy, việc con người vận hành con người dự đoán các vấn đề cũng như giám sát và giám sát hệ thống ngày càng trở nên khó khăn hơn. Các vấn đề có thể trở nên nghiêm trọng hơn nếu các hệ thống được đưa vào thực địa trước khi được kiểm tra đầy đủ và bởi các yếu tố cụ thể liên quan đến bối cảnh quân sự, chẳng hạn như tính bí mật của công nghệ và hệ thống hoặc việc sử dụng bầy đàn, trong đó lỗi của một đơn vị được chia sẻ với các đơn vị khác.

³⁷ Kalev Leetaru: 'Sẽ siêu âm Cuối cùng, vũ khí cũ ng đẩy chúng ta hướng tới phòng thủ tên lửa được hỗ trợ bởi AI?'. Forbes, ngày 23 tháng 6 năm 2019. <https://www.forbes.com/sites/kalevleetaru/2019/06/23/ý-chí-vũ-khí-siêu-thanh-cuối-cùng-đẩy-chúng-ta-về-phía-ai-cung-cấp-năng-lưu-lực-tên-lửa-phòng-thủ/#520c8ce04871>

Giảm thiểu tác động của AI

Các hệ thống vũ khí hoạt động với mức độ tự chủ khác nhau trong các chức năng quan trọng của chúng đã tồn tại và những tiến bộ mới không ngừng xuất hiện. Bất chấp những lợi thế quân sự mà các hệ thống hỗ trợ AI được cho là mang lại, như chúng ta đã thấy, vẫn có những rủi ro đáng kể liên quan đến việc sử dụng chúng.

Một báo cáo của Trung tâm Dân chủ Ash của Đại học Harvard Quản trị và Đổi mới về việc sử dụng nhân tạo một cách có đạo đức

thông tin tình báo trong các hoạt động của chính phủ cảnh báo rằng máy móc, ít nhất là vào thời điểm hiện tại, không được phép đưa ra những quyết định quan trọng liên quan đến cá nhân con người³⁸.

Báo cáo khuyến nghị rằng “AI chỉ nên được sử dụng để phân tích và cải tiến quy trình, chứ không nên hỗ trợ ra quyết định và sự giám sát của con người vẫn nên được áp dụng phổ biến”. Các hệ thống trí tuệ nhân tạo “không nên được giao nhiệm vụ đưa ra các quyết định quan trọng của chính phủ đối với công dân” - điều này rõ ràng sẽ bao gồm các quyết định của quân đội về sự sống hay cái chết của họ.

Nhiều biện pháp giảm thiểu khác nhau đã được đề xuất để Kiểm soát việc sử dụng AI trong các ứng dụng có rủi ro cao. Liên minh Châu Âu đã công bố dự thảo quy định về AI để đảm bảo an toàn và nhân quyền đồng thời cho phép tiếp thu công nghệ AI.³⁹ Dự thảo quy định tuân theo cách tiếp cận dựa trên rủi ro. Các hệ thống có rủi ro không thể chấp nhận được, được coi là mối đe dọa rõ ràng đối với sự an toàn, sinh kế và quyền lợi của người dân sẽ bị cấm. Cái này sẽ

bao gồm các hệ thống vũ khí tự động gây chết người. Các hệ thống AI có rủi ro cao, bao gồm tất cả các hệ thống nhận dạng sinh trắc học từ xa, sẽ phải tuân theo các nghĩa vụ nghiêm ngặt trước khi chúng có thể được đưa ra thị trường và sẽ được ghi lại trên cơ sở dữ liệu do Ủy ban Châu Âu duy trì. Các hệ thống có rủi ro hạn chế sẽ có nghĩa vụ minh bạch cụ thể để đảm bảo rằng người dùng hiểu rằng họ đang tương tác với máy.

Tổng thư ký Liên hợp quốc đã kêu gọi

LUẬT bị cấm theo luật pháp quốc tế,⁴⁰ và Ủy ban Chữ Thập Đỏ Quốc tế đã

khuyến nghị các quốc gia áp dụng các quy tắc ràng buộc pháp lý mới để cấm LUẬT và đảm bảo rằng vũ khí luôn nằm trong tầm kiểm soát của con người⁴¹. Những quy tắc như vậy phải áp dụng cho tất cả các hệ thống áp dụng vũ lực dựa trên việc xử lý đầu vào của cảm biến và phải đặt ra ngoài vòng pháp luật những hệ thống nhằm mục tiêu vào con người và các hệ thống mà con người không thể kiểm soát một cách có ý nghĩa, bao gồm cả những công nghệ không rõ ràng quá phức tạp để có thể hiểu được. Các hệ thống dựa trên cảm biến còn lại phải tuân theo các nghĩa vụ tích cực, chẳng hạn như về vị trí, thời gian sử dụng và thông số kỹ thuật mục tiêu của chúng, để bảo vệ luật chiến tranh hiện hành khỏi bị xói mòn.⁴²

³⁸ Modhana Ranindranath: 'Các cơ quan nên đề phòng AI phi đạo đức', Báo cáo cảnh báo'. Nextgov.com, ngày 23 tháng 8 năm 2017. <http://www.nextgov.com/cio-briefing/2017/08/cơ-quan-nên-xem-out-phi-đạo-đức-ai-report-cảnh-báo/140461>

³⁹ Ủy ban Châu Âu: 'Châu Âu phù hợp với Thời đại Kỹ thuật số: Ủy ban đề xuất các quy tắc và hành động mới để đạt được sự xuất sắc và tin tưởng vào Trí tuệ Nhân tạo'. Ngày 21 tháng 4 năm 2021. https://ec.europa.eu/hoa-hồng/góc-báo-chí-tiết/en/ip_21_1682

⁴⁰ 'Vũ khí tự động giết người phải bị cấm, người đứng đầu LHQ khẳng định'. Liên Hợp Quốc, ngày 25 tháng 3 năm 2019. <https://news.un.org/en/story/2019/03/1035381>

⁴¹ điểm của ICRC về hệ thống vũ khí tự động'. Ủy ban Chữ Thập Đỏ Quốc tế, ngày 12 tháng 5 năm 2021. <https://www.icrc.org/en/document/icrc-position-automatic-weapon-systems>

⁴² 'Quy định quyền tự chủ trong hệ thống vũ khí'. Điều 36, tháng 10 năm 2020. <https://article36.org/wp-content/uploads/2020/10/Điều-tiết-tự-chủ-leaflet.pdf>

Kết luận: dư ới sự kiểm soát của con người?



Các cường quốc quân sự lớn trên thế giới đang chạy đua giới thiệu công nghệ AI trước khi giải quyết những lo ngại nghiêm trọng nêu ra. Tín dụng: DoD

Bản tóm tắt này trình bày các ứng dụng quân sự khác nhau đã được hình dung cho AI, đồng thời nêu bật khả năng gây hại cho cá nhân của chúng. Như chúng ta đã thấy, các đề xuất nhằm giảm thiểu rủi ro do hệ thống AI quân sự gây ra dựa trên nguyên tắc đảm bảo rằng hệ thống AI luôn nằm dư ới sự giám sát của con người.

Như ng liệu con người có thể giữ quyền kiểm soát các hệ thống AI tiên tiến hay không? Trong cuộc trò chuyện cùng nhau, những người đoạt giải Nobel Kazuo Ishiguro và Venki Ramakrishnan dư ờng như không nghĩ vậy. Ishiguro cảnh báo rằng “Loài người còn tụt hậu quá xa nên không có cách nào bạn có thể theo dõi được con người. Thực sự, nó sẽ giống như việc có một người gác đêm đã nghỉ hưu u nào đó cố gắng giám sát một sân vận động đầy những người hâm mộ bóng đá náo loạn”, trong khi Ramakrishnan chỉ ra rằng “AI có thể đưa ra ý tứ ờng lớn tiếp theo, như dân chủ, chủ nghĩa cộng sản hoặc chủ nghĩa Quốc xã, hoặc tiền, hoặc công ty cổ phần. Một khi nó hiểu được cách điều khiển cảm xúc của con người, chúng ta sẽ có nhiều điều lớn hơn cần quan tâm”.⁴³

Khả năng gây hại từ AI vượt ra ngoài phạm vi cá nhân. Việc sử dụng rộng rãi AI có ý nghĩa xã hội sâu sắc và tiềm năng định hình lại thế giới một cách căn bản

⁴³ 'Kazuo Ishiguro và Venki Ramakrishnan: tư ờng tư ờng về một nhân loại mới'. Tài chính Times, ngày 26 tháng 3 năm 2021. <https://www.ft.com/content/eca7988d-2961-4b27-9368-ff58c966e969>

44 Jaron Lanier và Glen Weyl: 'AI là một
Tư tưởng, không phải là công nghệ'. Có đây,
ngày 15 tháng 3 năm 2020. [https://www.wired.com/
câu-chuyện/kiến-ai-is-an-ideology-not-a-
technology/](https://www.wired.com/câu-chuyện/kiến-ai-is-an-ideology-not-a-technology/)
45 Để biết thêm thông tin, vui lòng
xem <https://rethinkingsecurity.org.uk/>

của tư tưởng lai. Viết trên tạp chí 'Wired', Jaron Lanier và Glen Weyl nhận xét rằng "AI" được hiểu rõ nhất là một hệ tư tưởng chính trị và xã hội hơn là một tập hợp các thuật toán. Cốt lõi của hệ tư tưởng là một bộ công nghệ, được thiết kế bởi một nhóm nhỏ kỹ thuật tinh hoa, có thể và nên trở nên tự chủ và cuối cùng thay thế, thay vì bổ sung, không chỉ cho từng cá nhân mà còn cho phần lớn nhân loại." Họ chỉ ra rằng hệ tư tưởng này "có sự cộng hưởng mạnh mẽ với các hệ tư tưởng lịch sử khác, chẳng hạn như chế độ kỹ trị và các hình thức chủ nghĩa xã hội dựa trên kế hoạch hóa tập trung, được coi là mong muốn hoặc không thể tránh khỏi việc thay thế hầu hết sự phán xét/cơ quan của con người bằng các hệ thống được tạo ra bởi một tầng lớp kỹ thuật nhỏ".⁴⁴ Nhìn trong bối cảnh này, sự phát triển của AI quân sự chỉ có thể làm trầm trọng thêm những căng thẳng thực sự giữa chủ nghĩa quân phiệt và dân chủ.

Điều này sẽ đặt ra những câu hỏi nghiêm túc về cách thức và lý do AI được sử dụng, đặc biệt là trong các ứng dụng quân sự. Những câu trả lời không làm bạn yên tâm. Có nhiều rào cản về công nghệ, pháp lý và đạo đức cần phải vượt qua trước khi AI có thể được triển khai rộng rãi như các cường quốc trên thế giới, trong nỗ lực dẫn đầu về AI quân sự, đang chạy đua để đạt được mục tiêu đó.

giới thiệu công nghệ như vậy trước khi giải quyết những mối quan ngại này. Đọc những lời hoa mỹ trong Tạp chí Tích hợp và từ các nhà bình luận quân sự, khó tránh khỏi kết luận rằng phần lớn động lực cho sự quan tâm của Vương quốc Anh trong quân sự AI là phản ứng trước những diễn biến ở Mỹ, Nga, Trung Quốc: 'Nếu họ làm thì chúng ta không muốn bị bỏ lại phía sau'. Mặc dù vậy, những thách thức an ninh mà Vương quốc Anh phải đối mặt rất khác so với những thách thức mà các quốc gia khác phải đối mặt, đặt ra câu hỏi đầu là vấn đề mà chúng ta đang cố gắng giải quyết và có cách nào tốt hơn để giải quyết nó không? Drone Wars UK tin rằng việc sử dụng công nghệ tiên tiến không phải là cách duy nhất - hoặc thậm chí là tốt nhất - để giải quyết các mối lo ngại về an ninh và Vương quốc Anh nên hướng tới các phương pháp thay thế để duy trì an ninh quốc gia dựa trên việc bảo vệ và phúc lợi của người dân trung tâm của chính sách an ninh.⁴⁵

Cho đến nay, công chúng ít đánh giá cao những thay đổi và rủi ro mà xã hội đang phải đối mặt do những tiến bộ trong AI và robot. Bản tóm tắt này, theo cách riêng của nó, được coi là một phần của lời cảnh tỉnh. AI có thể và nên được sử dụng để cải thiện điều kiện tại nơi làm việc và dịch vụ cho công chúng, chứ không phải để làm tăng mức độ sát thương trong chiến tranh.

“Khả năng gây hại từ AI vượt ra ngoài phạm vi cá nhân. Việc sử dụng rộng rãi AI có ý nghĩa xã hội sâu sắc và tiềm năng định hình lại cơ bản thế giới trong tương lai”



Shining a spotlight
on military drones

Xuất bản tháng 12 năm 2021

Drone Wars Vũ dng quốc Anh, Ngôi nhà hòa bình,
19 Phố Paradise, Oxford, OX1 1LD

 01865 243688

 info@dronewars.net

 dronewars.net