سارة سكولز | نشر في ١٢ ديسمبر ٢٠٢٣,

البنتاغون وعسكرة الذكاء الاصطناعي

ترجمة وتحرير: نون بوست

رقم 4 هاملتون بليس هو مبنى ذو أعمدة يقع في وسط لندن، وهو مكان جمعية الطيران الملكية مع أربعة طوابق مخصصة للفعاليات. في شهر أيار/ مايو استضاف المنزل الإدواردي الذي يعود تاريخه إلى أوائل القرن العشرين اجتماعًا أكثر حداثة؛ حيث اجتمع مسؤولو الدفاع والمقاولون والأكاديميون من جميع أنحاء العالم لمناقشة مستقبل تكنولوجيا الطيران والفضاء العسكرية.

وسرعان ما سارت الأمور على نحو خاطئ. في ذلك المؤتمر؛ بدأ أن تاكر هاميلتون، رئيس اختبارات وعمليات الذكاء الاصطناعي في القوات الجوية للولايات المتحدة، يصف محاكاة مثيرة للقلق؛ حيث تم تكليف طائرة بدون طيار مدعمة بالذكاء الاصطناعي بإسقاط مواقع الصواريخ، ولكن عندما بدأ مشغل بشري بالتدخل في هذا الهدف، قتلت الطائرة بدون طيار مشغلها، وقطعت نظام الاتصالات.

وتلا ذلك الحماسة والخوف تجاه الإنترنت. ففي وقت يتزايد فيه القلق العام بشأن الذكاء الاصطناعي الجامح، يعتقد كثير من الناس، بما في ذلك المراسلين، أن القصة كانت حقيقية، لكن هاميلتون سرعان ما أوضح أن هذه المحاكاة التي تبدو بائسة لم يتم تنفيذها في الواقع، لقد كانت مجرد تجربة فكرية.

قالت إميليا بروباسكو، وهي زميلة بارزة في مركز الأمن والتكنولوجيا الناشئة بجامعة جورج تاون: "هناك الكثير مما يمكننا فهمه حول سبب انحراف هذه القصة".

جزء من السبب هو أن السيناريو قد لا يكون في الواقع بعيد المنال؛ فقد وصف هاميلتون قتل العميل بأنه "نتيجة معقولة" في تعليقاته اللاحقة. وتزداد أدوات الذكاء الاصطناعي قوة؛ حيث يقول بعض النقاد إن السيطرة عليها أصبحت أكثر صعوبة.

حريصة على التكنولوجيا

على الرغم من المخاوف بشأن أخلاقيات الذكاء الاصطناعي وسلامته؛ فإن الجيش يراهن بشكل كبير

على الذكاء الاصطناعي.، فلقد طلبت وزارة الدفاع الأمريكية 1.8 مليار دولار للذكاء الاصطناعي والتعلم الآلي في سنة 2024، بالإضافة إلى 1.4 مليار دولار لمبادرة محددة ستستخدم الذكاء الاصطناعي لربط المركبات وأجهزة الاستشعار والأشخاص المنتشرين في جميع أنحاء العالم. وقال بنجامين بودرو، الباحث السياسي في مؤسسة راند والمؤلف المشارك لتقرير بعنوان "التطبيقات العسكرية للذكاء الاصطناعي: مخاوف أخلاقية في عالم غامض": "لقد أبدت الولايات المتحدة اهتمامًا نشطًا للغاية بدمج الذكاء الاصطناعي في جميع وظائف الحرب".

في الواقع؛ فإن الجيش حريص للغاية على التكنولوجيا الجديدة لدرجة أن "المشهد العام هو نوع من الاستيلاء على الأراضي في الوقت الحالي لتحديد أنواع المشاريع التي ينبغي تمويلها"، كما كتب شون سميث، كبير المهندسين في شركة بلو هلو – وهي شركة مقاولات دفاعية تبيع الذكاء الاصطناعي والأنظمة المستقلة – في رسالة بالبريد الإلكتروني إلى أندارك. وتستثمر دول أخرى، بما في ذلك الصين، بكثافة في الذكاء الاصطناعي العسكري.

وفي حين أن قدرًا كبيرًا من القلق العام بشأن الذكاء الاصطناعي يدور حول آثاره المحتملة على الوظائف، فإن الأسئلة المتعلقة بالسلامة والأمن تصبح أكثر إلحاحًا عندما تكون حياة الناس على المحك.

ودفعت هذه الأسئلة إلى بذل جهود مبكرة لوضع حواجز أمام استخدام الذكاء الاصطناعي وتطويره في القوات المسلحة، في الداخل والخارج، قبل دمجه بالكامل في العمليات العسكرية. وكجزء من أمر تنفيذي صدر في أواخر تشرين الأول/ أكتوبر؛ كلف الرئيس جو بايدن بتطوير مذكرة للأمن القومي "سيضمن أن الجيش الأمريكي ومجتمع الاستخبارات يستخدم الذكاء الاصطناعي بشكل آمن وأخلاقي وفعال في مهامه، وسيوجه الإجراءات لمواجهة الاستخدام العسكري للخصوم للذكاء الاصطناعي".

بالنسبة لأصحاب المصلحة؛ تحتاج هذه الجهود إلى المضي قدمًا، حتى لو ظلت كيفية تنفيذها، وما هي تطبيقات الذكاء الاصطناعي المستقبلية التي سيتم تطبيقها عليها؛ غير مؤكدة.

وقال بودرو: "نحن نعلم أن أنظمة الذكاء الاصطناعي هذه هشة وغير موثوقة. وليس لدينا دائمًا تنبؤات جيدة حول التأثيرات التي ستنتج فعليًا عندما تكون في بيئات تشغيل معقدة".

ما الذي يتم استخدامه بالفعل؟

إن تطبيقات الذكاء الاصطناعي موجودة ومفعلة جيدًا داخل وزارة الدفاع، وهي تشمل الأمور العادية، مثل استخدام شات جي بي تي لتأليف مقال، ولكن المستقبل يحمل تطبيقات محتملة ذات أعلى قدر ممكن من الاهتمام، مثل أنظمة الأسلحة الفتاكة المستقلة: الأسلحة التي يمكنها التعرف على شخص ما وقتله بمفردها، دون موافقة بشرية على ضغط الزناد.

كما حرصت وزارة الدفاع أيضًا على دمج الذكاء الاصطناعي في مركباتها، حتى تتمكن الطائرات بدون طيار والدبابات من التنقل والتعرف على الأهداف وإطلاق الأسلحة بشكل أفضل؛ حيث تمتلك بعض الطائرات المقاتلة بالفعل أنظمة ذكاء اصطناعي تمنعها من الاصطدام بالأرض.

وقامت وكالة مشاريع الأبحاث الدفاعية المتقدمة "داربا" – إحدى منظمات البحث والتطوير التابعة لوزارة الدفاع – برعاية برنامج في سنة 2020 يضع طيارًا يعمل بالذكاء الاصطناعي في مواجهة طيار بشري. وفي خمس "معارك جوية" – معارك في الجو بين الطائرات المقاتلة – قاتل الطيار البشري بأسلوب "Gun Top" ضد الخوارزمية التي تحلق على نفس الطائرة المحاكاة. وبدا الإعداد وكأنه لعبة فيديو؛ حيث تظهر على الشاشة طائرتان نفاثتان تطاردان بعضهما البعض في السماء، ولقد فاز نظام الذكاء الاصطناعي؛ حيث أسقط الطائرات الرقمية التي يقودها الإنسان. وكان جزء من هدف البرنامج،

وفقًا لوكالة مشاريع البحوث المتطورة الدفاعية "داربا"، هو جعل الطيارين يثقون ويحترمون الروبوتات، وتمهيد الطريق لمزيد من التعاون بين الإنسان والآلة في السماء.

هذه كلها تطبيقات تقنية تتمحور حول الأجهزة، لكن وفقًا لبروباسكو، زميلة جامعة جورجتاون؛ فإن الرؤية الحاسوبية المدعمة بالبرمجيات – أي قدرة الذكاء الاصطناعي على استخلاص معلومات ذات معنى من الصورة – تعمل على تغيير الطريقة التي يتعامل بها الجيش مع البيانات المرئية. ويمكن استخدام مثل هذه التكنولوجيا، على سبيل المثال، لتحديد الأشياء في صورة التجسس عبر الأقمار الصناعية؛ حيث قالت برزباسكو: "يستغرق الأمر وقتًا طويلًا حقًا حتى ينظر الإنسان إلى كل شريحة ويكتشف شيئًا ما قد تغير أو شيئًا موجودًا".

يمكن لأنظمة الكمبيوتر الذكية تسريع عملية التفسير هذه، من خلال تحديد التغييرات (ظهرت الشاحنات العسكرية في مكان لم تكن فيه من قبل) أو أشياء محددة مثيرة للاهتمام (مثل طائرة مقاتلة)، حتى يتمكن الإنسان من إلقاء نظرة. وقالت بروباسكو، التي تبحث في كيفية إنشاء ذكاء اصطناعي جدير بالثقة ومسؤول لتطبيقات الأمن القومي: "يبدو الأمر كما لو أننا جعلنا الذكاء الاصطناعي يؤدي لعبة "أين والدو؟" للجيش".

وفي سياق مماثل؛ أجرت وحدة الابتكار الدفاعي – التي تساعد البنتاغون على الاستفادة من التقنيات التجارية – ثلاث جولات من مسابقة الرؤية الحاسوبية تسمى "إكس فيو"؛ حيث طلبت من الشركات إجراء تحليل آلي للصور المتعلقة بموضوعات تشمل الصيد غير القانوني أو الاستجابة للكوارث. ويمكن للمؤسسة العسكرية أن تتحدث علنًا عن هذا العمل الإنساني، وهو غير سري، وأن تشارك قدراتها ومعلوماتها مع العالم، لكن هذا لن يكون مثمرًا إلا إذا رأى العالم أن تطوير الذكاء الاصطناعي لديه قوي، فيما يحتاج الجيش إلى أن يتمتع بسمعة طيبة فيما يتعلق بالتكنولوجيا الصلبة. وقال ديفيد دانكس، أستاذ علوم البيانات والفلسفة بجامعة كاليفورنيا في سان دييغو: "يتعلق الأمر باحترام الناس خارج الولايات المتحدة لنا والاستماع إلينا والثقة بنا".

وسيكون هذا النوع من الثقة مهمًا بشكل خاص في ضوء تطبيق الذكاء الاصطناعي العسكري الشامل: برنامج يسمى القيادة والسيطرة المشتركة لجميع المجالات، والذي يهدف إلى دمج البيانات من جميع أنحاء القوات المسلحة. ففي جميع أنحاء الكوكب؛ تقوم الأجهزة الموجودة على السفن والأقمار الصناعية والطائرات والدبابات والشاحنات والطائرات بدون طيار وغيرها بإصدار معلومات باستمرار عن الإشارات المحيطة بها، سواء كانت تلك الإشارات مرئية أو صوتية أو تأتي في أشكال لا يستطيع البشر الشعور بها، مثل موجات الراديو. فبدًلا من عزل معلومات البحرية عن معلومات الجيش، على سبيل المثال، سيتم ربط معلوماتهم الاستخباراتية بعقل واحد كبير للسماح بالتنسيق وحماية الأصول؛ حيث أوضحت بروباسكو أنه "في الماضي إذا أردت إطلاق النار على شيء ما من سفينتي، كان عليّ أن أرصده بالرادار الخاص بي". إن برنامج القيادة والسيطرة المشتركة هو مشروع طويل الأمد لا يزال قيد التطوير، لكن الفكرة هي أنه بمجرد تشغيله، يمكن لأي شخص استخدام البيانات من رادار آخر (أو قمر صناعي) لتفعيل تلك القوة المميتة، وسيظل البشر ينظرون إلى تفسير الذكاء الاصطناعي، ويقررون ما إذا كانوا سيضغطون على الزناد (وهو تمييز مستقل قد لا يعني الكثير بالنسبة "للهدف").

نائبة الرئيس الأمريكي كامالا هاريس تلقي ملاحظات حول الذكاء الاصطناعي في سفارة الولايات المتحدة بلندن في الأول من تشرين الثاني/ نوفمبر.

"روبوتات دردشة" مع الأحياء

بدأت مسيرة بروباسكو المهنية بالفعل في البحرية؛ حيث خدمت على متن سفينة تستخدم نظام أسلحة يسمى إيجيس، وقالت: "في ذلك الوقت، لم نكن نطلق عليه اسم الذكاء الاصطناعي، ولكن الآن إذا نظرت إلى التعريفات المختلفة، فستجد أنه مؤهل". وقد ساهمت هذه الخبرة عندما كانت شابة في تشكيل عملها – الذي أخذها إلى البنتاغون وإلى مختبر الفيزياء التطبيقية بجامعة جونز هوبكنز – حيث إنها تضع في اعتبارها دائمًا الشباب الذين يبلغون من العمر 18 سنة، وهم يقفون أمام أنظمة التحكم في الأسلحة، ويحاولون لمعرفة كيفية استخدامها بشكل مسؤول، وقالت: "إنها أكثر من مجرد مشكلة فنية، ومشكلة تدريب، إنها مشكلة إنسانية".

ويمكن لبرنامج القيادة والسيطرة المشتركة أيضًا إنقاذ الأشخاص والأشياء، وهي حسابات أخلاقية تقوم بها وزارة الدفاع بالطبع. من الناحية النظرية، من شأنه أن يجعل الهجوم على أي سفينة حربية أمريكية (أو طائرة أو قمر صناعي) أقل تأثيرًا؛ حيث يمكن لأجهزة الاستشعار والأسلحة الأخرى أن تملأ المنصة المفقودة. على الجانب الآخر؛ قد يكون النظام الموحد أيضًا أكثر عرضة لمشاكل مثل الهجمات الإلكترونية: فمع تدفق جميع المعلومات عبر القنوات المتصلة، قد يؤدي القليل من الفساد إلى قطع شوط طويل.

تعتبر هذه الطموحات كبيرة – وهي طموحات يشكك بعض الخبراء في إمكانية تحقيقها ومن فائدتها – لكن الجيش يحب التعمق في البحث بشكل كبير على غرار موظفيه. ففي عالم تطوير الذكاء الاصطناعي العسكري الواسع والوحشي؛ يعد أحد الأشخاص الذين يعملون في مشروع للذكاء الاصطناعي لصالح

وزارة الدفاع هو عالم فيزياء نظرية يُدعى أليكس ألانيز، والذي يقضي معظم وقته خلف أسوار قاعدة كيرتلاند الجوية، في مختبر أبحاث القوات الجوية في ألبوكيرك، نيو مكسيكو.

في النهار؛ يعمل ألانيز كمدير برنامج مُحسّن مشاركة الأسلحة، وهو نظام ذكاء اصطناعي يمكنه مساعدة البشر على إدارة المعلومات في الصراع – مثل المسارات التي تسلكها الصواريخ، وحاليًا؛ يتم استخدامه للعب سيناريوهات افتراضية تسمى ألعاب الحرب. يستخدم الاستراتيجيون ألعاب الحرب للتنبؤ بكيفية تطور الموقف، لكن صاحب العمل في ألانيز أبدى مؤخرًا اهتمامًا بمشروع مختلف للذكاء الاصطناعي كان ألانيز يسعى إليه: ففي أوقات فراغه، صنع روبوت دردشة يحاول التحدث مثلما يفعل أينشتاين.

لصنع تقنية أينشتاين، حصل ألانيز على الكود البرمجي الذي يشغل برنامج الدردشة الآلي المتاح بشكل مفتوح عبر الإنترنت، وقام بتحسين أجزائه الداخلية، ثم صنع 101 نسخة منه. بعد ذلك، قام ببرمجة وتدريب كل نسخة لتكون خبيرة في جانب واحد من حياة أينشتاين، بحيث كان أحد الروبوتات يعرف التفاصيل الشخصية لأينشتاين، وآخر يعرف نظرياته النسبية، وآخر يعرف اهتمامه بالموسيقى، وأفكاره حول الحرب العالمية الثانية، ومن بين مهارات روبوتات الدردشة؛ قد يكون هناك تضاربات. ثم ربطها معًا، وأقام روابط بين مواضيع مختلفة وأجزاء من المعلومات. بعبارة أخرى؛ أنشأ شبكة رقمية يعتقد أنها يمكن أن تقود نحو الذكاء العام الاصطناعي، وأنواع الاتصالات التي يقوم بها البشر.

من الممكن تطبيق مثل هذه التكنولوجيا على البشر الأحياء أيضا: فإذا ابتكر الممارسون الحربيون محاكاة واقعية وتفاعلية لزعماء العالم – بوتين على سبيل المثال – فإن صراعاتهم المحاكية يمكن، من الناحية النظرية، أن تكون أكثر صدقًا مع الحياة وتساعد في عملية صنع القرار. على الرغم من أن هذه التكنولوجيا مملوكة حاليًا لشركة ألانيز، وليست تابعة لمختبر أبحاث القوات الجوية، ولا أحد يعرف تطبيقاتها النهائية، فإن المنظمة مهتمة باستعراضها، وهو أمر منطقي في مشهد الاستيلاء على الأراضي الذي وصفه سميث. وفي منشور على موقع لينكد إن للترويج لمحاضرة ألقاها ألانيز في مؤتمر إنسباير الذي عقده المختبر في تشرين الأول/أكتوبر، قالت المنظمة إن عمل ألانيز في أينشتاين "يدفع بأفكارنا حول الذكاء وكيف يمكن استخدامه للتعامل مع الأنظمة الرقمية بشكل أفضل".

لكن ما يريده حقًا هو تطوير آينشتاين إلى شيء يمكنه التفكير مثل الإنسان: الذكاء العام الاصطناعي المتبجح، النوع الذي يتعلم ويفكر مثل الإنسان البيولوجي. وهذا شيء لا يريده غالبية الأمريكيين بشكل قاطع، فوفقًا لاستطلاع حديث أجراه معهد سياسات الذكاء الاصطناعي؛ يعتقد 63 بالمائة من الأمريكيين أن "التنظيم ينبغي أن يهدف إلى تقييد الذكاء الاصطناعي بشكل فعال".

استخدم سميث، مهندس بو هالو، آينشتاين ورأى كيف يشغل عتاده، وهو يعتقد أن التكنولوجيا من المرجح أن تصبح نوعًا من المساعدة لنماذج اللغات الكبيرة – مثل روبوت الدردشة شات جي بي تي – الذي شاع استخدامه مؤخرًا. ويمكن لمجموعة المعلومات الأكثر تقييدًا والتحكم التي يمكن لروبوت مثل أينشتاين الوصول إليها أن تحدد حدود استجابات نموذج اللغات الكبيرة، وتوفر ضوابط على إجاباته. وقال سميث إن أنظمة التدريب وحواجز الحماية هذه يمكن أن تساعد في منعهم من إرسال "استجابات وهمية غير دقيقة".

مخاوف بشأن مخاطر الذكاء الاصطناعي

من بين المخاوف المتعلقة بتكنولوجيا الذكاء الاصطناعي التأكد من أن النماذج اللغوية الكبيرة لا "تهذي" بمعلومات خاطئة. لكن المخاوف تمتد إلى سيناريوهات أكثر خطورة، بما في ذلك نهاية العالم الناجمة عن الذكاء الاصطناعي المارق. مؤخرًا، وقع كبار العلماء وخبراء التكنولوجيا على بيان جاء فيه ببساطة أن "التخفيف من خطر الانقراض الناجم عن الذكاء الاصطناعي ينبغي أن يكون أولوية عالمية إلى جانب المخاطر الأخرى على نطاق مجتمعي مثل الأوبئة والحرب النووية".

قبل ذلك بوقت قصير وتحديدا في شهر آذار/ مارس 2023، وقعت شخصيات بارزة أخرى – بما في ذلك المؤسس المشارك لشركة آبل ستيف وزنياك، وماكس تيغمارك من معهد الذكاء الاصطناعي للذكاء الاصطناعي والتفاعلات الأساسية في معهد ماساتشوستس للتكنولوجيا – خطابًا مفتوحًا بعنوان "إيقاف تجارب الذكاء الاصطناعي العملاقة مؤقتًا". وقد دعت الوثيقة إلى القيام بذلك لتحديد ما إذا كانت تأثيرات تجارب الذكاء الاصطناعي العملاقة إيجابية ويمكن التحكم في مخاطرها. واقترح الموقعون أن يأخذ مجتمع الذكاء الاصطناعي فترة توقف لمدة ستة أشهر "للتطوير والتنفيذ المشترك لمجموعة من بروتوكولات السلامة المشتركة لتصميم وتطوير الذكاء الاصطناعي المتقدم يتم تدقيقها والإشراف عليها بدقة من قبل خبراء خارجيين مستقلين".

اتفق الموقعون على أن "المجتمع سبق له التوقف مؤقتًا عن استخدام تقنيات أخرى ذات آثار كارثية محتملة على المجتمع"، مستشهدين بمساعي مثل الاستنساخ البشري، وتحسين النسل، وهم يعتقدون أنه "يمكننا أن نفعل ذلك في هذه الحالة". لكن لم يحدث هذا التوقف، ووصف النقاد مثل إميلي بندر من جامعة واشنطن الرسالة نفسها بأنها مثال على الصخب حول الذكاء الاصطناعي – في إشارة إلى الميل إلى المبالغة في تقدير القدرات الحقيقية للأنظمة وقربها من الإدراك البشري.

قال بودرو، الباحث في مؤسسة راند، إنه مع ذلك ينبغي على قادة الحكومات أن القلق بشأن ردود الفعل السلبية. وأضاف "هناك الكثير من الأسئلة والمخاوف حول دور الذكاء الاصطناعي، والأهم من ذلك أن تظهر الولايات المتحدة التزامها بالاستخدام المسؤول لهذه التقنيات". قد تُغذّي فكرة قيام طائرة بدون طيار متمردة بقتل مشغلها الكوابيس البائسة. على الرغم من أن هذا السيناريو قد يكون شيئًا يريد البشر تجنبه بوضوح، إلا أن لجنة الأمن القومية المختصة في الذكاء الاصطناعي يمكن أن تقدم مآزق أخلاقية أكثر دقة – لاسيما أن الجيش يعمل وفقًا لمجموعة مختلفة من القواعد.

فعلى سبيل المثال، قال نيل رو، أستاذ علوم الحاسوب في كلية الدراسات العليا البحرية "هناك استخدامات مثل جمع المعلومات الاستخبارية، التي تعتبر سلوكا مقبولاً من قبل الدول". ويمكن لأي دولة أن تراقب سفارة داخل حدودها وفقًا لاتفاقية فيينا لسنة 1961، لكن شركات التكنولوجيا ملزمة بمجموعة مختلفة تمامًا من اللوائح. لا يستطيع المساعد الذكي الذي يعرف باسم أليكسا أمازون تسجيل وتخزين المحادثات التي يسمعها دون إذن مسبق – على الأقل إذا كانت الشركة لا ترغب في توريط نفسها في مشاكل قانونية. وعندما يقتل جندي مقاتلاً في صراع عسكري، فإن ذلك يعتبر إصابة وليس جريمة قتل، حسب قانون الصراع المسلح. تسمح الاستثناءات المنصوص عليها في ميثاق الأمم المتحدة قانونيا باستخدام القوة المميتة في ظروف معينة.

في السياق ذاته، قال دانكس، الأستاذ في جامعة كاليفورنيا سان دييغو، إن "الأخلاقيات تدور حول قيمنا وكيف ندرك تلك القيم من خلال أفعالنا الفردية والجماعية. في عالم الأمن القومي، يميل الناس إلى الاتفاق حول قيم – السلامة والأمن والخصوصية واحترام حقوق الإنسان – ولكن "هناك بعض الخلاف حول كيفية التوفيق بين الصراعات والقيم". يعتبر دانكس من بين مجموعة صغيرة من الباحثين – بما في ذلك الفلاسفة وعلماء الحاسوب – الذين أمضوا سنوات في استكشاف تلك الصراعات في قطاع الأمن القومي ومحاولة إيجاد الحلول المحتملة.

انخرط لأول مرة في أخلاقيات الذكاء الاصطناعي والعمل السياسي جزئيًا لأن الخوارزمية التي ساعد في تطويرها لفهم بنية الدماغ البشري من المحتمل أن تستخدم داخل مجتمع الذكي، لأنه على حد تعبيره تشبه الاتصالات الداخلية للدماغ تلك الموجودة بين عقد خلية إرهابية. لكن ليس هذا ما صممناه من أجله على الإطلاق". وأضاف: "أشعر بالذنب لأنني لم أتفطن لذلك من قبل. ولكن كان ذلك بمثابة صحوة لأنني صرت أعرف كيف أريد استخدام الخوارزمية التي أطورها، لا أستطيع أن أفترض أن هذه هي

الطريقة الوحيدة التي يمكن استخدامها بها. إذن ما هي واجباتي كباحث؟".

بدأ دانكس عمله في هذا المجال بالتركيز على المجال العسكري، وبعد ذلك عندما شاهد مدى تأثير الذكاء الاصطناعي في تشكيل الحياة اليومية، عممه على مجالات أخرى على غرار السيارات ذاتية القيادة. واليوم، تميل أبحاثه إلى التركيز على قضايا مثل التحيّز في الخوارزميات، والشفافية في "تفكير" الذكاء الاصطناعي. لكن، بالطبع، لا يوجد ذكاء اصطناعي واحد ولا نوع واحد. تقدم خوارزميات التعلم الآلي التي تقوم بالتعرف على الأنماط أو التنبؤ بها وطرق الاستدلال المنطقي ببساطة إجابات بنعم أو لا على ضوء مجموعة من الحقائق. ويؤكد رو أن الجيوش فشلت إلى حد كبير في فصل الشكل الذي يبدو عليه الاستخدام الأخلاقي للأنواع المختلفة من الخوارزميات، وهي فجوة تناولها في ورقة بحثية نشرت سنة 2022 في مجلة "فرونتيرز إن بيغ داتا".

وكتب رو أن أحد الاعتبارات الأخلاقية المهمة التي تختلف بين أنظمة الذكاء الاصطناعي هو ما إذا كان النظام قادرًا على تفسير قراراته. فعلى سبيل المثال، إذا قرر الحاسوب أن هذه الشاحنة معادية، فيجب أن يكون قادرًا على تحديد الجوانب التي أدت إلى هذا الاستنتاج. حيال ذلك، قال رو "عادةً، إذا كان التفسير جيدًا، فيجب أن يكون الجميع قادرين على فهمه".

جندي من القوات الجوية الأمريكية و"كلب آلي" أثناء تمرين بالقرب من لاس فيغاس.

الخوارزميات العسكرية

فيما يتعلق بنوع الرؤية الحاسوبية في مسلسل الأطفال التلفزيوني"أين والدو" التي ذكرها بروباسكو، يجب أن يكون النظام قادرًا على تفسير سبب تصنيف شاحنة على أنها عسكرية أو طائرة نفاثة على أنها "مقاتلة". يمكن للإنسان بعد ذلك مراجعة العلامة وتحديد ما إذا كانت كافية أم معيبة. فعلى سبيل المثال، بالنسبة للشاحنة التي تم تحديدها على أنها مركبة عسكرية، قد يعلل الذكاء الاصطناعي اختياره من خلال الإشارة إلى اللون والحجم المطابقين لها. لكن عند مراجعة هذا التفسير، قد يرى الإنسان أنه

فشل في النظر إلى لوحة الترخيص أو رقم التسلسل أو شعارات الشركة المصنعة التي من شأنها أن تكشف أنه نموذج تجاري.

لا يمكن للشبكات العصبية، التي تُستخدم غالبًا للتعرف على الوجه أو غيرها من مهام الرؤية الحاسوبية، تقديم هذه التفسيرات بشكل جيد. ويقول رو إن عدم القدرة على تفسير استجابتها يشير إلى أن هذا النوع من الخوارزميات لن يساهم في بناء ذكاء عام اصطناعي – النوع الشبيه بالإنسان – لأن البشر يمكنهم عادة توفير الأساس المنطقي. لأسباب أخلاقية ونظرا للحدود المحتملة لعموميتها – على حد تعبير رو – يمكن أن تشكل الشبكات العصبية مشكلة بالنسبة للجيش، الذي شرع إلى جانب القوات الجوية والبحرية بالفعل في تنفيذ مشاريع الشبكات العصبية.

أما بالنسبة لأنواع الخوارزميات التي يمكنها تفسير نفسها، يقترح رو أن يبرمجها الجيش لعرض الحسابات المتعلقة بمخاوف أخلاقية محددة، مثل عدد الضحايا المدنيين المتوقع من عملية معينة. ويمكن للبشر بعد ذلك مسح تلك المتغيّرات وتقييمها والانتقاء منها. فعلى سبيل المثال، ستؤثّر كيفيّة ترجيح المتغيرّات على طياري الذكاء الاصطناعي، كما هو الحال في السيارات ذاتية القيادة. وإذا كان على أنظمة تجنّب الاصطدام الاختيار بين إصابة طائرة عسكرية من جانب الطائرة وعلى متنها إنسان، أو إصابة طائرة خاصة بها مدنيون أجانب، فيحتاج المبرمجون إلى التفكير في ماهية القرار "الصحيح" في هذه الحالة مسبقًا.

وعلى نحو مماثل، يتعيّن على مطوري الذكاء الاصطناعي في القطاع التجاري أن يطرحوا أسئلة مماثلة حول المركبات ذاتية القيادة. لكن بشكل عام، يتعامل الجيش مع مسألة الحياة والموت أكثر من الشركات الخاصة. وأوضح دانكس أنه "في السياق العسكري، أعتقد أنه من الشائع للغاية أن نستخدم التكنولوجيا لمساعدتنا في اتخاذ خيارات مستحيلة".

بطبيعة الحال، إن الانخراط في العمليات التي تؤدي إلى هذه المقايضات على الإطلاق حساب أخلاقي في حد ذاته. وكما هو الحال مع المثال أعلاه، تؤكّد بعض الأطر الأخلاقية أنه يجب إبقاء البشر على اطلاع ليتخذوا القرار النهائي بشأن قرارات الحياة أو الموت على الأقل. لكن حتى هذا – كما يقول رو – لا يجعل استخدام الذكاء الاصطناعي أكثر أخلاقية بشكل مباشر. قد يكون لدى البشر معلومات أو هواجس أخلاقية لا تمتلكها الروبوتات، لكن يمكن للروبوتات تجميع المزيد من البيانات، لذلك قد تفهم أشياء لا يفهمها البشر، فضلا عن أن البشر أيضًا ليسوا موضوعيين أو قد يكونون غير أخلاقيين بشكل فعال.

قد يكون للذكاء الاصطناعي أيضًا تحيزاته الخاصة. تستشهد دراسة رو بنظام افتراضي يستخدم البيانات الأمريكية لتحديد القوى "الصديقة" بصريًا. في هذه الحالة، قد تجد أن المباريات الودية تميل إلى أن تكون طويلة، وبالتالي يتم وضع علامة على الأشخاص قصار القامة كأعداء، أو تفشل في وضع علامة على الخصوم الطويلين على هذا النحو. وقد يؤدي ذلك إلى وفاة – أو استمرار حياة – أشخاص غير مقصودين، خاصة إذا كانت أنظمة الأسلحة الفتاكة المستقلة هي التي تصدر الأوامر. في هذه الحالة، أصبح هناك أمران واضحان: لدى الجيش مصلحة نشطة في جعل الذكاء الاصطناعي الخاص به أكثر موضوعية، والدافع ليس مجرد السعي وراء طيبة القلب: بل جعل الهجوم والدفاع أكثر دقة. وقال دانكس إن "الشيء الأكثر أخلاقية يكمن دائمًا تقريبًا في عدم الدخول في حرب، أليس كذلك؟ والمشكلة هي ماذا نفعل حيال ذلك؟ بما أنه انتهى بنا المطاف إلى هذا الوضع".

ما هي المعايير الأخلاقية؟

هناك جهود لتحديد المبادئ الأخلاقية التي ستلتزم بها وزارة الدفاع عندما يتعلق الأمر بالذكاء الاصطناعي. في سنة 2019، اقترح مجلس الابتكار الدفاعي، وهو منظمة استشارية مستقلة تابعة لوزارة الدفاع تقدم المشورة العسكرية جزئيًا بشأن التكنولوجيا الناشئة، مجموعة من مبادئ الذكاء

الاصطناعي. من أجل إعداد تلك المبادئ، عقدوا سلسلة من جلسات الاستماع العامة – بما في ذلك في جامعة كارنيجي ميلون، حيث كان دانكس رئيسًا للقسم في ذلك الوقت. وأوضح قائلا: "كان هناك انفتاح حقيقي على فكرة أنه لم يكن لديهم بالضرورة كل المعرفة، أو حتى ربما الكثير من المعرفة، عندما يتعلق الأمر بقضايا الأخلاق والمسؤولية والجدارة بالثقة وما شابه ذلك".

بعد التوصيات التي خرج بها مجلس الإدارة وأبحاثه، أصدرت وزارة الدفاع مبادئ أخلاقيات الذكاء الاصطناعي في سنة 2020، بما في ذلك أن تكون الأنظمة "قابلة للتتبع"، مما يعني أن الموظفين سيفهمون كيفية عمل الأنظمة، وسيكون لديهم إمكانية الوصول إلى "منهجيات شفافة وقابلة للتدقيق، ومصادر البيانات وإجراءات التصميم والوثائق". كما تنص المبادئ على أن الذكاء الاصطناعي يجب أن يكون قابلاً للتحكم، مما يعني أنه يمكن للإنسان إيقاف تشغيله إذا أصبح مارقًا (وحاول قتل مشغله مثلا). وقال بودرو إن "هناك الكثير من الشيطنة في التفاصيل، ولكن هذه هي المبادئ التي إذا تم تنفيذها بنجاح، من شأنها أن تخفف قدرا كبيرا من المخاطر المرتبطة بالذكاء الاصطناعي".

وفي الوقت الحالي، تعد المبادئ التوجيهية الشاملة الخاصة بالذكاء الاصطناعي لوزارة الدفاع ككل مجرد كلام: على الرغم من أنه يجب على المتعاقدين العسكريين الالتزام بالعديد من السياسات والممارسات العامة فيما يتعلق بكل ما يبيعونه، إلا أن هناك مجموعة محددة من السياسات أو الممارسات التي يفتقر مطورو الذكاء الاصطناعي العسكري إلى الالتزام بها، كما هو الحال مع الاتفاق حول كيفية استخدام الذكاء الاصطناعي في الحرب.

يشير بودرو إلى أن وزارة الدفاع توظف محامين للتأكد من امتثال الولايات المتحدة للقانون الدولي، مثل اتفاقيات جنيف. ولكن بعيدًا عن الحماية الإنسانية العالمية من هذا القبيل، لا يوجد قانون دولي قابل للتنفيذ خاص بالذكاء الاصطناعي. لقد قام الناتو، مثل الجيش الأمريكي، ببساطة بوضع إرشادات في شكل "مبادئ الاستخدام المسؤول للذكاء الاصطناعي في الدفاع". وتطلب الوثيقة من دول الناتو وحلفائه النظر في مبادئ مثل الشرعية، وقابلية التفسير، وقابلية الحكم، وتخفيف التحيز "كخط أساس للحلفاء عند استخدام الذكاء الاصطناعي في سياق الدفاع والأمن".

وأكد دانكس أنه من المهم الاتفاق، على المستوى الدولي، ليس فقط على المبادئ وإنما أيضا على كيفية تحويلها إلى ممارسات وسياسات. وهو يعتقد أن "بعض هذه المناقشات قد بدأت. لكنها مبكرة للغاية. ومن المؤكد، على حد علمي، أنه لم يتم التوصل إلى اتفاقات بشأن المعايير المتفق عليها والقابلة للتطبيق دوليا". وفي سنة 2022، وقعت كاثلين هيكس، نائبة وزير الدفاع الأمريكي، وثيقة تسمى استراتيجية الذكاء الاصطناعي المسؤول ومسار التنفيذ، وهي خطوة واحدة نحو تحويل المبادئ إلى عملية في الولايات المتحدة. وأوضح دانكس قائلا "أعتقد أنها تترجم، ولكن فقط من خلال العمل".

في وحدة الابتكار الدفاعي، أراد المسؤولون محاولة تحويل المبادئ إلى براغماتية. وبما أن هذا هو عمل وزارة الدفاع، فإن ذلك يتضمن أوراق عمل للمشاريع، ويطرح أسئلة مثل "هل حُددت المهام بوضوح؟" وهم يرون ذلك وسيلة محتملة للتخفيف من مشاكل مثل التحيز. خذ مثلا التحيز العنصري في التعرف على الوجه، كما يقول برايس غودمان، كبير الاستراتيجيين للذكاء الاصطناعي والتعلم الآلي في وزارة الدفاع. وأضاف أن السبب في ذلك غالبًا هو أن المطورين لم يكلفوا أنفسهم بجعل خوارزمياتهم دقيقة عبر المجموعات، بدلاً من توفير الدقة الشاملة. وفي مجموعة بيانات تضم 95 جنديًا أبيض وخمسة جنود سود، مثلا، يمكن لنظام الذكاء الاصطناعي أن يكون دقيقًا من الناحية الفنية بنسبة 95 بالمئة في التعرف على الوجوه – بينما يفشل في التعرف على 100 بالمئة من الوجوه الجنود السود.

وقال غودمان إن "الكثير من هذه الأنواع من الألغاز الأخلاقية التي تستحوذ على العناوين الرئيسية هي أخطاء نراها تتدفق من نوع من إدارة المشروع غير المتقنة والوصف الفني غير المتقن".

على المستوى الدولي، تهتم البلاد بالتوصل إلى توافق في الآراء بشأن أخلاقيات الذكاء الاصطناعي العسكري. وفي مؤتمر انعقد في لاهاي في شباط/ فبراير، أطلقت الحكومة الأمريكية مبادرة تهدف إلى تعزيز التعاون الدولي في مجال الذكاء الاصطناعي، وأنظمة الأسلحة المستقلة، والمسؤولية بين جيوش العالم. (أشارت نسخة محفوظة من الإطار إلى أن البشر يجب أن يحتفظوا بالسيطرة على الأسلحة النووية، وليس تسليم الأزرار إلى الروبوتات، ولم يظهر هذا البيان في الإطار النهائي).

لكن إقناع الدول غير الحليفة بالموافقة على الحد من التكنولوجيا التي تمتلكها، وبالتالي ربما قوتها، ليس بالأمر السهل. من الممكن أن تشجع المنافسة الولايات المتحدة على تسريع وتيرة التنمية ـ وربما تجاوز الضمانات على طول الطريق. ومن جهته، لم يستجب مكتب وكيل وزارة الدفاع للأبحاث والهندسة في وزارة الدفاع لطلب التعليق على هذه النقطة.

حسب بودرو "هدف هذه الضغوط التأكد من أن الولايات المتحدة ليست في وضع تنافسي غير مؤاتٍ لأنها تتصرف بطريقة أكثر تقييدًا من خصومها". وإحدى نسخ هذه الديناميكية هي الطريقة التي بدأت بها الولايات المتحدة العمل على الأسلحة النووية لأول مرة: اعتقد العلماء أن ألمانيا كانت تعمل على تصنيع قنبلة ذرية، وبالتالي يجب على الولايات المتحدة أن تبدأ أيضًا، حتى تتمكن من تطوير مثل هذا السلاح أولاً وتكون لها الأسبقية. واليوم، لا تزال ألمانيا لا تمتلك أي أسلحة نووية خاصة بها، في حين تمتلك الولايات المتحدة الآلاف منها. فهم ما يمكن أن يحدث في المستقبل، خاصة من دون حواجز الحماية الصحيحة، يعدّ محور التدريبات الفكرية مثل تصور طائرة مسيّرة تقتل مشغلها. وقال بروباسكو إنه "من الصعب حقًا معرفة ما هو قادم وكيفية المضي قدمًا".

المصدر: وورلد كرانش

رابط المقال: https://www.noonpost.com/185622/