

أخلاقيات الذكاء الاصطناعي... تغوُّل الآلة ومسؤولية الإنسان

ضوابط لكبح جماح البرامج والنظم

على الرغم من أن العالم بدأ التعرف على مصطلح «الذكاء الاصطناعي» خلال مؤتمر لكلية دارتموث الأميركية في خمسينات القرن الماضي؛ فإن القضية الأخلاقية بشأن تطوراتها احتاجت لعقود حتى تتبلور وتتطور، إذ تناسبت طردياً -ولا تزال- مع نمو تطبيقاته وانتشارها منذ المحاولات البسيطة الأولى حتى اللحظة التي يتفاعل فيها البشر راهناً «تشات جي بي تي» ونسخه المتلاحقة.

بالنسبة إلى العاملين في قطاع التكنولوجيا سواء على مستوى الصناعة أو الابتكار فإن ثمة وجهين لقضية «الأخلاقيات»؛ فمن جهة يدعو فريق لإخضاع التطورات المرتقبة للمسألة الأخلاقية قبل تطبيقها وعدّها معياراً حاكماً إلى حد المطالبة بـ«هدنة لحكومة التطبيقات، لضمان عدم إضرارها بالبشر»، غير أن ذلك على الوجه الآخر «قد يعطل أو يدفع المطورين للارتداد إلى الخلف»، وفق ما يقدر فريق آخر.

وأجمع «خبراء وأكاديميون» تحدثوا إلى «الشرق الأوسط» على «أهمية وحيوية تطبيقات الذكاء الاصطناعي، ومدى الرفاهية والفائدة التي أضافتها لحياة الكثيرين»، غير أنهم يدعون مع ذلك إلى ضرورة وضع ضوابط أخلاقية «تضبط جماح البرامج التي تتصرف آلياً، ولكن من دون». «أن تتحول تلك المعايير إلى سيف على رقبة الذكاء الاصطناعي».

وكان صناع تكنولوجيا ومختصون في الذكاء الاصطناعي، قد وقفوا في مارس (آذار) الماضي عريضة تطالب بـ«وقف أبحاث تطوير الذكاء الاصطناعي ستة أشهر لإتاحة الفرصة نحو مزيد من الحكمة لهذا». «النشاط، لضمان عدم تضرر البشر منه».

وأثارت تلك العريضة التي أعدها معهد «فيوتشر أوف لايف» غير الربحي، واقترب عدد الموقعين عليها من 3 آلاف عالم ورائد أعمال، مخاوف من أن «السباق الذي تخوضه مختبرات الذكاء الاصطناعي لتطوير ونشر عقول رقمية أكثر قوة، قد يخرج عن سيطرتها، بحيث لا يمكن لأحد،». «ولا حتى لمنشئوها، فهمها أو التحكم فيها بشكل موثوق».

توصيات أممية

ورغم ما يحققه الذكاء الاصطناعي من مزايا للبشرية، في تشخيص الأمراض والتنبؤ بتغيرات المناخ وغيرها من الفوائد، فإن تحدياته الأخلاقية، المتمثلة في بعض الخوارزميات المتحيزة جنسياً وعرقياً، والتهديد المحتمل لبعض تطبيقاته للخصوصية، دفعت منظمة الأمم المتحدة للتربية والعلوم والثقافة (يونسكو) إلى تبني «أول اتفاق عالمي بشأن توصيات حول أخلاقيات الذكاء الاصطناعي» في نوفمبر (تشرين الثاني) من عام 2021، وتتمثل في 4 توصيات، أولها: «حماية البيانات» بحيث تكون هناك إجراءات لحماية بيانات الأفراد وحقوقهم في التحكم بها. والثانية: «حظر استخدام نظم الذكاء الاصطناعي لأغراض المراقبة الجماعية» ووجوب اقتصار المسؤولية النهائية في أي مهمة على البشر بحيث لا تصبح تقنيات الذكاء الاصطناعي بحد ذاتها شخصية معنوية. أما الثالثة فتتمثل في «دعم وسائل الذكاء الاصطناعي التي تتسم بالكفاءة في استخدام البيانات والطاقة والموارد» لكي تسهم هذه الوسائل في التصدي لمعالجة القضايا البيئية. وجاءت التوصية الرابعة لتنص على أن تكون هناك «آلية لتقييم العواقب الأخلاقية للذكاء الاصطناعي».

دعوات متواصلة

المخاوف من الذكاء الاصطناعي «قديمة قدم المفهوم نفسه، ولكن روبوت الدردشة (تشات جي بي تي)، قرّبها كثيراً من نطاق الخيال العلمي إلى الواقع العملي» وفق ما يقول خوسيه ديلغادو، المتخصص في علم النفس التجريبي بجامعة غرناطة بإسبانيا، في تصريحات لـ«الشرق الأوسط».

الأدباء والمفكرون والفلاسفة وكذلك العلماء المتخصصون استشرفوا أهمية وجود أخلاقيات للذكاء الاصطناعي، إذ تحدث الروائي الأميركي فيرنور فينغ عام 1983، عن أن «الوجود المستقبلي للبشرية قد يعتمد على تنفيذ معايير أخلاقية راسخة في نظم الذكاء الاصطناعي نظراً لأن «هذه النظم، في مرحلة ما، قد تتطابق أو تحل محل القدرات البشرية».

وحذّر الفيلسوف السويدي نيك بوستروم في عام 2018 من المخاطر المحتملة للتميز التكنولوجي في حالة تحويل الآلات الذكية ضد «مبدعيها، أي البشر، وشدد على بناء «ذكاء صناعي ودود».

وتوافق العلماء مع المفكرين في التحذير من المخاوف الأخلاقية، فتحدثت عالمة الحاسوب الأميركية البارزة، روزاليند بيكار في عام 1997، عن أنه «كلما زادت حرية الآلة، احتاجت إلى معايير أخلاقية». «أكثر

وبالعودة إلى ديلغادو، فإنه لا يُبدي شكاً في أن روبوت الدردشة «تشات جي بي تي» سينجح في تجاوز أخطائه، بل إنه يعتقد أن «التعلم الآلي قد يقوده إلى أن يكون أذكى من الإنسان، وهنا تكمن الخطورة». «عند توظيف الذكاء الاصطناعي في عمل مشترك مع الإنسان

ولذلك فإن ديلغادو ينبه بشدة إلى ضرورة أن «يكون للإنسان السيطرة الكاملة والمسؤولية عن سلوك ونتائج نظم الذكاء الاصطناعي». مرجعاً ذلك إلى أنه «عند تحليل بعض الحوادث التي وقعت في السنوات الأخيرة، وجد الباحثون أن سببها تغول الذكاء الاصطناعي، ومنها (تحطم قطار ألفيا على الطريق من مدريد إلى سانتياغو دي كومبوستيلا في إسبانيا عام 2013، وتحطم رحلة الخطوط الجوية الفرنسية 447 عام 2009)، وتحطم رحلة الخطوط الجوية آسيانا 214 عام 2013

ويضيف: «خلص الباحثون الذين درسوا هذه الحوادث إلى أن سببها الأساسي يكمن في أن استراتيجيات التحكم باستخدام الذكاء الاصطناعي لم تكن مماثلة أو مشابهة لتلك المستخدمة من المراقبين البشريين

وهذا التحدي الذي حدده ديلغادو في التفاعل بين البشر والذكاء الاصطناعي، خلاصته أنه «لتعزيز علاقة أخلاقية وعادلة بين البشر ونظم الذكاء الاصطناعي، من الواجب أن تستند التفاعلات إلى المبدأ الأساسي «لا احترام القدرات المعرفية للبشر

الإنسان الاستثنائي

ولا يتعارض التحدي السابق مع تحدّي ثانٍ، حدده أوزليم غاريباي، الأستاذ المساعد في قسم الهندسة الاصطناعية ونظم الإدارة بجامعة كاليفورنيا الأميركية، وهو «الذكاء الاصطناعي المسؤول» الذي يعني برأيه أن «يكون داعماً لرفاهية الإنسان، ولكن بطريقة تتماشى مع «القيم البشرية

ويقول غاريباي لـ«الشرق الأوسط»: «في إطار هذا التحدي، يمكن أن تقدم الروبوتات الذكية حلاً طبيعياً لبعض أشكال الإعاقة، لكن لا ينبغي أن يتطور الأمر إلى (تأليه التقنية) واستخدامها لبناء

(الإنسان الاستثنائي أو الفائق) عبر تحسين مواصفاته وتعزيز ذاكرته
«مثلاً باستخدام شرائح إلكترونية».

بعُد أو تحدٍ آخر لفت إليه مارك أنطوان ديلاك، الأستاذ المساعد
المتخصص في الأخلاقيات والفلسفة السياسية في جامعة مونتريال بكندا،
في مقال نشره بمجلة «رسالة اليونيسكو» التابعة للمنظمة الأممية في
مارس 2018، لافتاً إلى «البرمجيات التي يتم تطبيقها بالفعل في
الكثير من البلدان لتحديد «السلوك الإرهابي» أو «الشخصية
الإجرامية» لدى الأفراد، باستخدام تقنية التعرف على ملامح الوجه.
وقال إن باحثين من جامعة «ستانفورد» بالولايات المتحدة تعرضوا
للذعر «من هذا الانبعاث الجديد لنظرية الفراسة التي تحلل شخصية
«الفرد اعتماداً على معالم وجهه وتعابير».

لكن الفراسة المنحازة لم تكن أمنية وقائية فقط، بل إن تطبيقات
التوظيف تمارس جانباً من هذا، إذ كشفت دراسة لدورية «إم آي تي
تكنولوجي ريفيو» (يُصدرها معهد «ماساتشوستس للتكنولوجيا») في
فبراير (شباط) 2022 أن مواقع مثل «لينكد إن» تسعى «لإزالة بعض
برمجياتها الخاصة بمقابلات العمل، والتي كانت تتحيز ضد أصحاب
«الهمم، والنساء المرشحات للعمل».

استخدام الأسلحة

ويقول يوشوا بنجيو، وهو عالم كندي من أبرز متخصصي الحاسوب
المعاصرين، وحاصل على جائزة «تورنغ» 2018 (يُنظر إليها بوصفها
موازية لجائزة «نوبل» في علوم الحاسب) لـ«الشرق الأوسط»، إنه يجب
العمل على «منع تصميم نظم الذكاء الاصطناعي التي تنطوي على مخاطر
«عالية للغاية مثل النظم التي يمكنها استخدام الأسلحة».

ويضيف بنجيو أن «نظم الذكاء الاصطناعي يمكن أن تحقق فوائد جمّة
لل بشرية، كما يمكن لتطبيقاتها أن تكون مفيدة في تحقيق الرعاية
الصحية، لكن من جهة أخرى يمكن تطوير نظم تستخدم الأسلحة، وهذا ما
«يجب العمل على حظره».

ومن التحديات الأخرى التي يجب التعامل معها الحرص على «تأمين
الخصوصية»، بحيث لا تنتهكها نظم الذكاء الاصطناعي، وهو أحد الأسباب
التي دعت إيطاليا إلى حظر «تشات جي بي تي»، وانضمت إليها إسبانيا
وفرنسا.

وأعلن «مجلس حماية البيانات الأوروبي»، عن إنشاء فريق لتعزيز تبادل المعلومات بشأن أي إجراءات يمكن اتخاذها تجاه «تشات جي بي تي»، وأفاد المجلس في 14 أبريل (نيسان) الماضي بأنه يؤيد «التقنيات المبتكرة في الذكاء الاصطناعي»، لكنه شدد على أنها «يجب أن تكون دائماً متوافقة مع حقوق الناس وحررياتهم».

و«يجمع تطبيق (تشات جي بي تي) البيانات الشخصية ويعالجها لتدريب خوارزمياته، وهذا انتهاك واضح للخصوصية»، كما يقول دومينيكو تاليا، أستاذ هندسة الكمبيوتر بجامعة «كالابريا» الإيطالية، «لـ«الشرق الأوسط».

ويضيف تاليا: «أنا مع هذا التطبيق وما يقدمه من مزايا لحياة البشر، لكن في الوقت نفسه لا أقبل أن يتم جمع بياناتي الشخصية عند «تفاعلي معه».

معاهدات دولية

التحديات السابقة التي يفرضها الذكاء الاصطناعي، يرى بنجيو، أنه «يجب التعامل معها في إطار قوانين وتشريعات ملزمة، وليس عبر التنظيم الذاتي»، ويضيف: «الأمر أشبه بالقيادة، سواء كان ذلك على الجانب الأيسر أو الأيمن، حيث يجب على الجميع القيادة بنفس الطريقة، وإلا سنكون في ورطة».

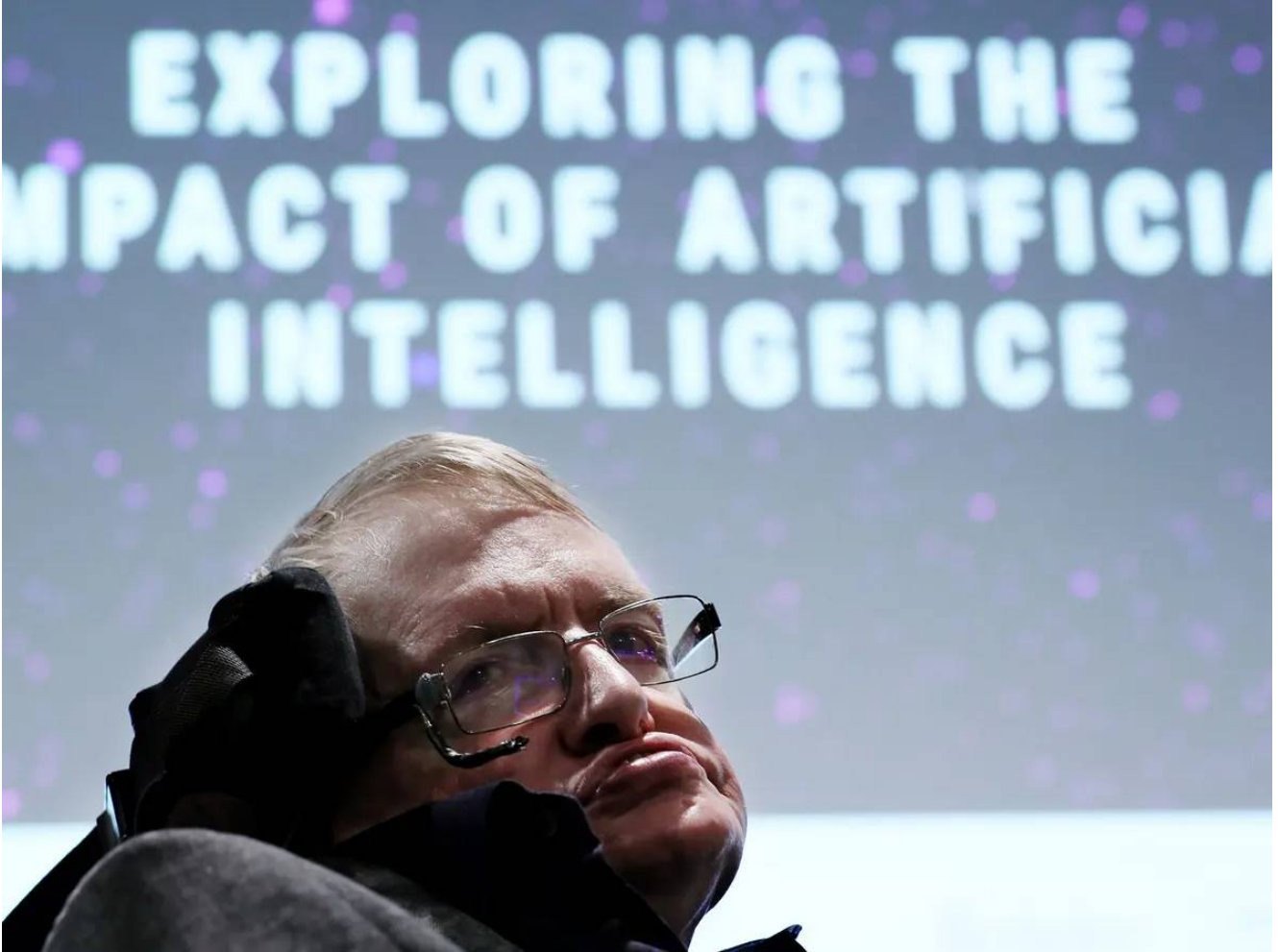
كما يلفت إلى أن «هناك مشروع قانون (بشأن الذكاء الاصطناعي) قيد الإعداد في الاتحاد الأوروبي، كما سيتم إقرار قانون قريباً في كندا، ولكن ذلك لا يُغني عن إصدار معاهدات دولية مماثلة لما حدث مع «(المخاطر النووية)، و(الاستنساخ البشري».

وبينما أصدرت «يونسكو» توصياتها بشأن التحديات الأخلاقية للذكاء الاصطناعي قبل أقل من عامين، فإن هناك أكثر من 40 دولة تعمل مع المنظمة لتطوير «ضوابط وتوازنات الذكاء الاصطناعي على المستوى الوطني».

ودعت المنظمة الأممية جميع البلدان للانضمام إلى الحركة التي تقودها لبناء «ذكاء صناعي أخلاقي»، مشيرةً في بيان نشرته في 30 مارس الماضي، إلى أنه «سيتم تقديم تقرير مرحلي عملاً تحقق في هذا الإطار خلال (منتدى اليونسكو العالمي) حول أخلاقيات الذكاء الاصطناعي بسلوفينيا في ديسمبر (كانون الأول) المقبل».

مركز «ليفريوم»... نداء ستيفن هوكينغ لأنسنة الذكاء الاصطناعي

أسسه الفيزيائي البارز عام 2016



الفيزيائي الراحل ستيفن هوكينغ خلال حفل افتتاح «ليفريوم» عام (2016) (جامعة كامبردج)
من المرجح أن يكون الذكاء الاصطناعي أفضل أو أسوأ شيء يحدث»
للشريحة على الإطلاق، لذلك فهناك قيمة كبيرة في تصحيحه»، عبر هذه
النداء، وقبل وفاته بعامين، حاول العالم البريطاني البارز ستيفن
هوكينغ (1942: 2018) تلخيص دعوته إلى «أنسنة الذكاء الاصطناعي».
وإفادة هوكينغ التي جاءت خلال افتتاح «مركز ليفريوم لمستقبل
الذكاء الاصطناعي» بجامعة كامبردج البريطانية (عام 2016) تعكس
جانباً من عمل المركز الذي يتخصص في «مستقبل الذكاء الاصطناعي»،
ومع ذلك فإنه يحرص على إظهار البعد الإنساني في اهتماماته وأوراقه
البحثية الصادرة عنه.

وتجمع الأبحاث التي يصدرها «ليفريوم» بين أكاديميين من تخصصات متنوعة؛ مثل: «التعلم الآلي، والفلسفة، والتاريخ، والهندسة، وغيرها»، وذلك بهدف «استكشاف إمكانات الذكاء الاصطناعي على المدى القصير والطويل، وضمان استخدامه لصالح البشرية»، وهو الهدف الذي حدده البروفسور هوكينغ، عند افتتاح المركز الذي تم تمويل إنشائه بمنحة 10 ملايين جنيه إسترليني.

وكذلك قال ستيفن كيف، مدير المركز في حفل افتتاحه، إننا «بحاجة إلى ذكاء صناعي يساعدنا على اتخاذ قرارات أفضل، لا أن يحل محلنا»، مضيفاً أنه يجب العمل على «ضمان أن تكون للنظم الاصطناعية الذكية أهداف تتماشى مع القيم الإنسانية، والتأكد من أن أجهزة الكمبيوتر لا تتطور تلقائياً في اتجاهات جديدة غير مرحب بها».

وخلال الأبحاث الأحدث لـ«ليفريوم»، سعى المتخصصون إلى «رصد التحيز الجنسي للرجل عن طريق إنشاء صورة نمطية ثقافية حول سيطرة الرجال على مجال الذكاء الصناعي»، الأمر الذي قد يؤدي، وفق تقديرهم، إلى «مزيد من النقص في عدد النساء العاملات بالمجال، فضلاً عن تسرب التحيز إلى الخوارزميات الموضوعة مثلاً لاختيار العاملين الجدد في أي مؤسسة».

وخلال الدراسة المنشورة في 13 فبراير (شباط) الماضي بدورية حصر باحثو «ليفريوم» 142 «Public Understanding of Science»، عملاً سينمائياً خلال قرن (1920: 2020) تطرقت للذكاء الاصطناعي، وحددوا 116 شخصية ظهرت كمحترفة بالمجال، وكان من بينهم 92 في المائة من الرجال، بينما تقدر النسبة الفعلية للرجال في القطاع بـ78 في المائة. وأبدى الباحثون مخاوفهم من أن الخيال العلمي يشكل الواقع، وقد يساعد «ترسيخ هذا الوضع عبر السينما، إلى تهميش المرأة في منتجات الذكاء الاصطناعي».

كما تمكن مشروع بحثي آخر أطلقه المركز خلال الفترة من 2018 إلى 2022 تحت عنوان «السرديات العالمية للذكاء الاصطناعي»، من رصد الروايات الخيالية للذكاء الاصطناعي وتحديد القيم والمصالح التي توجهها، وتحليل تأثيرها على الخيال العام والقبول العام، وعلى صانعي السياسات والحكومات.

القاهرة - حازم بدر

المصدر: صحيفة الشرق الأوسط