نشر في العدد 37 من
دورية اتجاهات الأحداث

تصميم غير عادل

تحيزات الذكاء الاصطناعي وتكريس اللامساواة

إلياس معوض

باحث في الابتكار الحكومي والاستشراف الاستراتيجي، لبنان.

أصبح التقدم السريع في إنشاء التقنيات التكنولوجية المختلفة قوة جبارة -لا يمكن إنكارها- تشارك في تشكيل ملامح المجتمع. فلم تعد جميع الابتكارات من الذكاء الاصطناعي إلى التعلم الآلي والحوسبة الكمومية والتكنولوجيا الحيوية، مجرد أدوات عادية؛ بل أصبحت عوامل فاعلة للتغيير الجوهري، أعادت صياغة الطريقة التي نعيش، ونعمل، ونتواصل، ونتعلم، ونصل إلى الأخبار والمعلومات بها.

كذلك لم تعد هذه التقنيات التكنولوجية الناشئة مقتصرة على مجال مختبرات الأبحاث أو الصناعات المتخصصة كما في السابق؛ بل أصبحت متغلغلةً في كل جوانب حياتنا، حتى صارت بمثابة محفزات للتغيير المجتمعي العميق، معيدةً بذلك تشكيل آليات السلطة، والهياكل الاقتصادية، والأعراف والسلوكيات الثقافية. أضف إلى هذا أن وتيرة تطور هذه التقنيات كانت مذهلة؛ بحيث بات كل سبقٍ تكنولوجي يقربنا أكثر إلى مستقبل كان حتى وقت قريب مجرد خيال علمي.

في الواقع إن هذه التقنيات تحمل أملاً بالتوصل إلى درجات من الكفاءة والإبداع والتقدم غير المسبوقين، بالإضافة إلى أنها تقدم حلولاً لبعض التحديات الأكثر إلحاحاً والتي تواجه البشرية الآن. فعلى سبيل المثال، تعمل هذه التقنيات التكنولوجية الحديثة في مجال الاقتصاد على زيادة المكاسب الإنتاجية، وتعزيز الابتكار، وإنشاء فرص جديدة لريادة الأعمال وخلق (أنواع جديدة من) الوظائف. وفي مجال التعليم، فإن هذه التقنيات توفر أدوات تحويلية للتعلم الشخصي، وتنمية المهارات، والوصول إلى المعرفة على نطاق واسع. وفي مجال الرعاية الصحية، تستهدف تعزيز دقة التشخيص، وتحسين نتائج العلاج، ودعم وتطوير الأدوية، وتمكين المرضى من أخذ زمام المبادرة فيما يتعلق بصحتهم. وفي القطاع العام؛ تتيح تلك التقنيات الفرصة لصنع القرارات المستندة إلى البيانات، وتعزز تقديم الخدمات العامة، وترفع مستوى المسؤولية والشفافية.

على الرغم من كل الفوائد الممكنة لهذه لتكنولوجيات الناشئة، فإن هناك اعترافاً متزايداً بالآثار السلبية والتعقيدات الخطرة التي يمكن أن تتسبب بها في بعض الأحيان. فمثلاً هناك الكثير والكثير من الأدلة التي تظهر كل يوم حول الدور والتأثير السلبيين لخوارزميات الذكاء الاصطناعي، وخاصةً فيما يتعلق بقدرتها على زيادة عدم المساواة والأضرار الاجتماعية، والتي يُعد أبرزها استنساخ القوالب النمطية الجنسانية والعرقية.

في هذه المقالة، نتعمق في تغطية دور وآثار التقنيات التكنولوجية الناشئة في زيادة وتحفيز عدم المساواة الاجتماعية، وسنستقصي أيضاً الأسباب الكامنة وراء المشكلة، ونعرض أمثلة من الحياة الواقعية، وندرس آخر التطورات في هذا المجال، ونختتم المقالة بمجموعة من الأسئلة والاعتبارات للمضي قدماً. ونهدف من خلال دراستنا للتفاعل المعقد بين الابتكار التكنولوجي والديناميكيات المجتمعية، إلى اكتساب فهم أعمق للتحديات والفرص التي تنتظرنا من أجل مستقبل أكثر شمولاً وإنصافاً.

> هناك أدلة على التأثير السلبي لخوارزميات الذكاء الاصطناعي، خاصة فيما يتعلق بعدم المساواة والأضرار الاجتماعية، والتي يُعد أبرزها استنساخ القوالب النمطية الجنسانية والعرقية.

غير عادلة في تكوينها

لكي نفهم كيف لهذه التقنيات أن تولد وتزيد من عدم المساواة الاجتماعية؛ فمن المهم أن نعرف من الذي يقوم بتطويرها، وكيف يتم هذا التطوير، ولأي غرض، وفي أي سياق. لقد أصبح من المعترف به والموثق على نطاق واسع الآن هو أن الأنظمة الآلية، المصممة لتكون موضوعية ومحايدة، غالباً ما تقوم بالتمييز بطرق تعكس التحيزات البشرية[1]؛ ومن ثم فإنها تُبقي على، أو حتى تفاقم التفاوتات الاجتماعية وأوجه عدم المساواة.

ويرجع ذلك إلى سببين رئيسيين: السبب الأول، هو أن سمات وآليات التنمية الأساسية لهذه الأنظمة الخوارزمية تكرر الأوجه المتأصلة من عدم المساواة، والسبب الثاني والأهم:

هـو الاعتمـاد المفـرط عـلى القـدرات التقنيـة في صنـع القـرار، وفصـل صنـع القـرار تمامـاً عـن سـياقاته الاجتماعيـة والتاريخيـة.

لفهـم كيفيـة تأثـير الأنظمـة الخوارزميـة عـلى حياتنـا اليوميـة بشـكل أفضـل، مـن المهـم أن نفهـم كيفيـة بنـاء هـذه الأنظمـة، وكيـف تعمـل، وكيـف يمكـن أن تـؤدي التصميمـات المتأصلـة فيهـا إلى إعـادة إنتـاج الأضرار. يوفـر التعلـم الآلي الخوارزميـات والتقنيـات الأساسـية التي تسـمح لأنظمـة صنـع القـرار -المدعومـة بالـذكاء الاصطناعـي- بتحليـل البيانـات، وتحديـد الأنمـاط، وإجـراء التنبـؤات، وصنـع القـرارات بأقـل قـدر مـن التدخـل البـشري.

وتتضمـن عمليـة التعلـم الآلي عـدة خطـوات؛ أولاً، جمـع البيانـات ذات الصلـة، والتـي يتـم بعـد ذلـك تنظيفهـا وإعدادهـا للتحليـل. بعـد ذلـك، يتـم اختيـار نمـوذج -يسـتخدم القواعـد والتعليمـات الرياضيـة- مناسـب وتقديمـه للتعلـم الآلي وتدريبـه باسـتخدام البيانـات ذات الصلـة؛ ممـا يسـمح للنمـوذج بتحديـد الأنمـاط والعلاقـات. في الخطـوة التاليـة يتـم تقييـم أداء النمـوذج باسـتخدام بيانـات اختبـار منفصلـة؛ ومـن ثـم تُجـرى تعديـلات لرفـع مسـتوى دقتـه. بمجـرد صقـل النمـوذج، فـإن بإمكانـه التنبـؤ بالأحـداث واتخـاذ القـرارات بنـاءً عـلى بيانـات جديـدة. وأخـيراً، يتـم اسـتعمال النمـوذج في تطبيقـات العالـم الحقيقـي وتتـم مراقبتـه وتحديثـه باسـتمرار للحفـاظ عـلى فعاليتـه.

وبالرغـم مـن اتبـاع كل هـذه الخطـوات في دورة حيـاة التعلـم الآلي، فـإن سـوريش وجوتـاج[2] قـد قامـا برصـد سـبعة أضرار عـبر هـذه الـدورة، بـدءاً مـن مظاهـر التحيـزات **التاريخيـة** أثنـاء جمـع البيانـات. ويحـدث هـذا عندمـا تعكـس البيانـات التحيـزات المجتمعيـة القائمـة أو مظاهـر عـدم المسـاواة (عـلى سـبيل المثـال: اسـتخدام الموافقـة المتحيزة عـلى القـروض أو بيانات السـجن) التـي تديـم الممارسـات التمييزية/العنصريـة. ويمكـن للتحيـزات **التمثيليـة** أيضـاً أن تنشـأ أثنـاء جمـع البيانـات حـين لا تكـون عينـة التدريـب عاكسـة بدقـة لسـكان العالـم الحقيقيـين؛ ممـا سـينتج عنـه نقـص التمثيـل أو إنشـاء عينـات معيبـة.

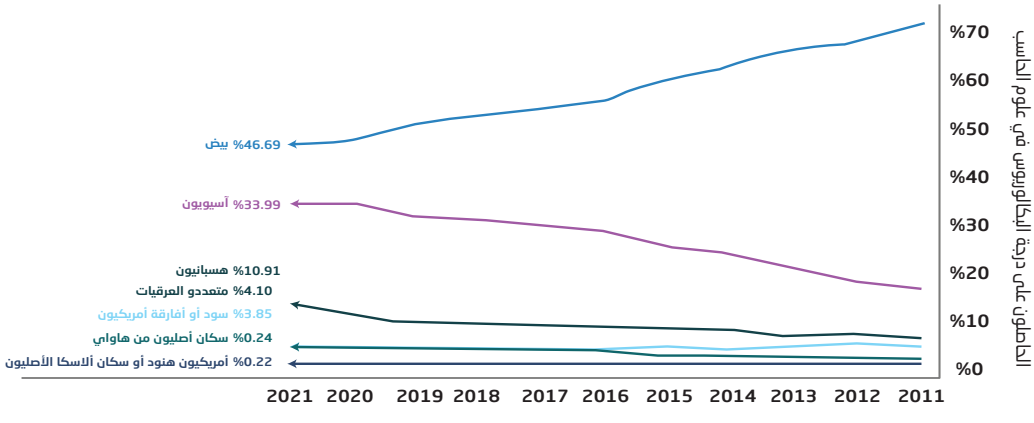
ويظهـر تحيـز **القيـاس** في مرحلـة إعـداد البيانـات عندمـا يتـم اختيـار صفـات أو محـددات بشـكل غـير دقيـق أو قياسـها بشـكل غـير متسـق عـبر المجموعـات، بينمـا يظهـر التحيـز **التجميعـي** أثنـاء تطويـر النمـوذج عندمـا يُعمـم تطبيـق نمـوذج واحـد عـلى مجموعـات متنوعـة، عـلى فرضيـة وجـود علاقـات موحـدة بـين كل المدخـلات والمخرجـات؛ ممـا يـؤدي إلى خفـض مسـتوى أداء بعـض المجموعـات دون غيرهـا. ويـبرز تحيـز **التعليـم** في كيفيـة تدريـب النمـوذج؛ إذ يمكـن أن تـؤدي الاختيـارات التـي تُتخـذ إلى تضخيـم اختلافـات الأداء بـين المجموعـات. أمـا تحيـز **التقييـم**، فيحـدث أثنـاء مرحلـة تقييـم النمـوذج؛ إذا كانـت مجموعـة البيانـات المرجعيـة لا تمثـل سـكان العالـم الحقيقـي؛ ممـا قـد يـؤدي إلى ضعـف الأداء الفعـلي في الواقـع. وأخـيراً، ينشـأ تحيـز **الاسـتخدام** عنـد عـدم تطابـق الغـرض المقصـود مـن النمـوذج واسـتخدامه العمـلي، وغالبـاً مـا يكـون ذلـك بسـبب الأنظمـة المعقـدة التـي تتضمـن صنـاع قـرار مـن البـشر ممـن يفـسرون مخرجـات النمـوذج بطـرق غـير صحيحـة.

عندما تكون البيانات المعبرة عن الأنماط المجتمعية القائمة على أرض الواقع هي ما يتم إدخاله لتدريب هذه الأنظمة؛ فإن القرارات الآلية الناتجة ستعكس عدم المساواة الاجتماعية القائم وتبقي عليه وترسخه[3] [4]؛ ومن هنا فإن التعلم الآلي سيعيد إنتاج "حلقات ردود الفعل"[5] التي تُكرر وتُضخم الأنماط المجتمعية القائمة[6]. ومن ثم فمن المنصف أن نقول إن الخوارزميات في جوهرها هي مجرد امتداد لمطوريها، ولذا فهي ليست محايدة على الإطلاق. بعبارة أخرى إن الخوارزميات تُشكل المجتمع بنفس الطريقة التي شُكلت بها وتم تدريبها عليها؛ ولهذا السبب فمن المهم تجاوز "المنتج" (التقنيات التكنولوجية الحديثة) وفحص الملفات الشخصية لمطوريه.

المطورون الحقيقيون للأنظمة الخوارزمية

يتم تدريب الأنظمة الخوارزمية باستخدام قواعد البيانات التي أنشأها البشر، لذا، فإنها تعكس بطبيعتها التحيزات البشرية والتسلسلات الهرمية للسلطة، وتعمل على زيادة عدم المساواة في التمثيل والنتائج. ويمكن فهم هذه الطبيعة المتحيزة بشكل أفضل من خلال التمعن في طبيعة مجموعات الباحثين الذين يعملون على تطوير هذه الأنظمة، ويسلط تقرير مؤشر الذكاء الاصطناعي لعام 2023 الصادر عن جامعة ستانفورد[7] الضوء بوضوح على الخلل العرقي الكبير بين أعداد خريجي مجال علوم الكمبيوتر من بيض وآسيويين في أمريكا؛ إذ يُظهر أن الأقليات العرقية الثلاث الأقل تمثيلاً، لا تأتي فقط من المجتمعات الضعيفة، ولكنها أيضاً تُعد الأكثر خضوعاً للتحيز الآلي، كما هو مبين في الرسم البياني 1.

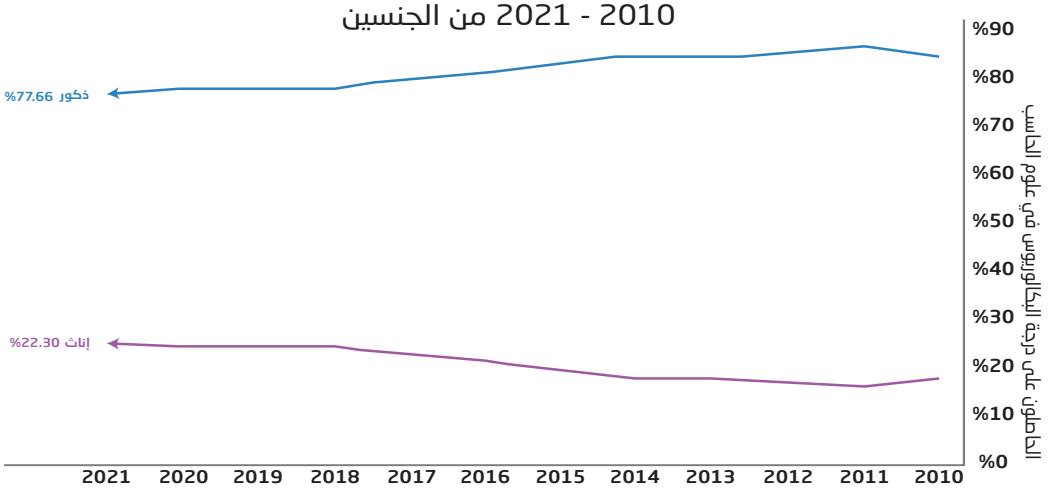
شكل (1) نسب عرقيات طلاب علوم الحاسوب المتخرجين في أمريكا الشمالية
في الفترة 2011 - 2021



Source: AI Index,Stanford University, 2023, https://aiindex.stanford.edu/ai-index-report-2023/

ويؤثر الشيء نفسه في تمثيل الجنسين؛ إذ تشكل النساء أقل من ربع طلاب وخريجي جامعات علوم الكمبيوتر، كما هو موضح في الرسم البياني 2.

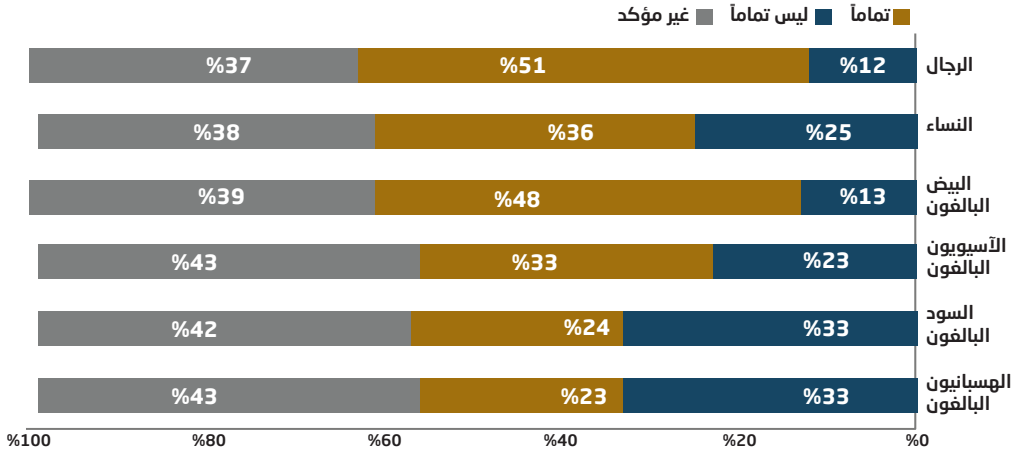
شكل (2) خريجو علوم الحاسب الجدد في أمريكا الشمالية
2010 - 2021 من الجنسين



Source: AI Index,Stanford University, 2023, https://aiindex.stanford.edu/ai-index-report-2023/

علاوة على ذلك، فإن الرسم البياني أدناه -الذي يسلط الضوء على المجموعات التي يتم أخذ تجاربها ووجهات نظرها في الاعتبار عند تطوير النماذج الخوارزمية- يظهر بوضوح أن وجهات النظر والخبرات السائدة التي يتم إدخالها عند تصميم هذه الأنظمة تنتمي في معظمها إلى رجال بالغين من البيض.

شكل (3) الأشخاص الذين تم تضمين آرائهم وخبراتهم في نظم الذكاء الاصطناعي
(النسب المؤية الإجمالية) في 2022



Source: AI Index,Stanford University, 2023, https://aiindex.stanford.edu/ai-index-report-2023/

إن إعادة إنتاج عدم المساواة لا تعتمد فقط على التقنية التكنولوجية المستخدمة، بل تعتمد أيضاً على عملية تعليم وتدريب تلك التقنية. ففي حين يمكن افتراض -ولو

بشكل خاطئ- أن التقنيات التكنولوجية محايدة؛ فإن البشر الذين يطورونها ليسوا كذلك بالتأكيد.

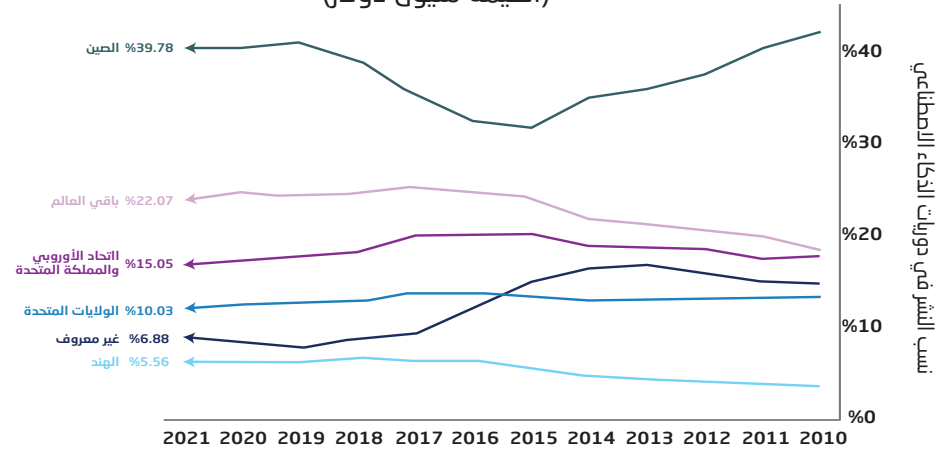
لقد قمنا حتى الآن بتغطية أوجه عدم المساواة التي أعادت التقنيات التكنولوجية الحديثة (الخوارزميات، والأشخاص الذين يقفون خلفها) إنتاجها. ومع ذلك يكشف نظام الذكاء الاصطناعي الأوسع نطاقاً عن طبقات ومستويات أخرى من عدم المساواة والأضرار المتفاقمة، والتي تتعلق على وجه التحديد بقدرات الإنتاج، ونماذج الأعمال، والتأثير البيئي.

إنتاج وتطوير الذكاء الاصطناعي

تعتمد تكلفة تطوير وتدريب الذكاء الاصطناعي على خمسة عناصر هي: البحث والتطوير (الخاص بالموظفين)، ورقائق رفع سرعة الذكاء الاصطناعي، والربط البيني على مستوى المجموعة، ومكونات الخادم، وتكاليف الطاقة. وتقتصر قدرة امتلاك معظم هذه العناصر إلى حد كبير على البلدان الأكثر ثراءً والأكثر تقدماً حول العالم.

وكما هو موضح بالرسوم البيانية أدناه، فإن قدرات البحث والتطوير، التي تمثل بداية خط إنتاج الذكاء الاصطناعي، تتركز بشدة في البلدان الأكثر ثراءً وتقدماً كما ذكر آنفاً.

شكل (4) النشر في الدوريات العلمية حول الذكاء الاصطناعي وفقاً للمناطق الجغرافية
(القيمة مليون دولار)



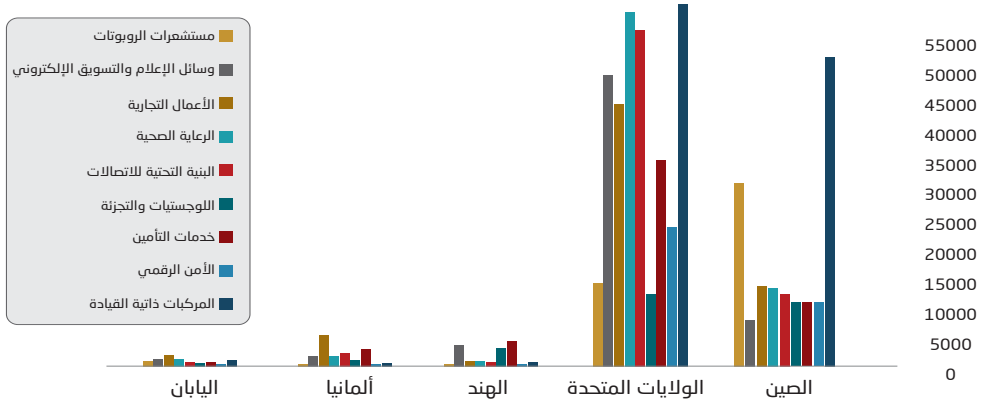
Source: AI Index,Stanford University, 2023, https://aiindex.stanford.edu/ai-index-report-2023/

وفي حين أنه من المفترض أن البلدان الأقل تقدماً قادرة على تنمية البحث والتطوير، فإنه يتعين عليها أن تتعامل مع تكلفة تدريب نماذج الذكاء الاصطناعي، والتي نادراً ما تكشف الشركات عنها. ولكن تشير التقديرات إلى أن تكلفة تدريب هذه الأنظمة أصبحت أكثر ارتفاعاً اليوم خاصةً مع نمو حجمها. ويقدر الخبراء أن هذه التكلفة كانت تنمو

بمعدل ضعفين إلى ثلاثة أضعاف سنوياً على مدى السنوات الثماني الماضية؛ مما يشير إلى أن النماذج الكبرى للذكاء الاصطناعي يمكن أن تتكلف أكثر من مليار دولار للنموذج بحلول عام 2027[8]. وهكذا فبمثل هذا المعدل من ارتفاع التكلفة، فمن المتوقع أن عدداً قليلاً فقط من المؤسسات سيتمكن من مواكبة هذه النفقات؛ مما سيحد من قدرة مشاركة الجميع في الابتكار والتطوير، كما أنه سيحتفظ بهذه القدرات في داخل حدود بعض الدول فقط[9].

وفقاً لتقرير مؤشر الذكاء الاصطناعي لعام 2023، بلغت الاستثمارات الخاصة عالمياً في مجال الذكاء الاصطناعي 91.9 مليار دولار عام 2022؛ وهو ما يزيد 18 مرة عما كانت عليه عام 2013 [10]، مع هيمنة الولايات المتحدة والصين على المشهد.

شكل (5) رأس المال الاستثماري في الذكاء الاصطناعي عبر قطاعات مختلفة 2018 - 2023



Source: AI Index,Stanford University, 2023, https://aiindex.stanford.edu/ai-index-report-2023/

إن تكلفة إنتاج الذكاء الاصطناعي تجعل من المستحيل عملياً على البلدان النامية وشركاتها المشاركة في هذا المجال، ناهيك عن المنافسة فيه، نظراً لهيمنة الفاعلين العالميين الحاليين.

يوضح كل من جوشي ونيكفيست الآتي: "يبدو من الصعب تحقيق جوانب العدالة فيما يتعلق بمجال الذكاء الاصطناعي عندما نجد أن التطورات في هذا المجال مرتبطة بشدة بالقدرة على الوصول -غير المتاح للجميع- إلى قواعد بيانات عملاقة وعلى توفير البنية التحتية الحاسوبية اللازمة لمعالجة هذه البيانات". والأكثر إثارة للقلق هو تركز الثروة والتنمية الناتجين عن تلك الأنظمة غير العادلة في أماكن بعينها حيث "يتحكم عدد صغير من الدول والشركات في تطوير الذكاء الاصطناعي وتطبيقاته؛ مما يثير أسئلةً مهمة فيما يتعلق بالانفصال الجغرافي بين أماكن تطوير الذكاء الاصطناعي والأماكن التي تحصد مكافآت الناتج المحلي الإجمالي"[11].

بالإضافة إلى الاحتكار المتزايد للذكاء الاصطناعي، تأتي مخاطر حماية الملكية الحصرية له، والتحكم في المواد الرقمية بالاعتماد على حقوق ملكية فكرية عفا عليها الزمن.

ومما يزيد الطين بلة أن تكلفة الذكاء الاصطناعي ليست متعلقة فقط بالتكاليف المادية الباهظة التي يحتاجها تطويره وتدريب نماذجه، والتي يمتلكها فقط الأثرياء والأقوياء، بل إن له أيضاً تكلفة أخرى كبيرة ألا وهي التكلفة البيئية، ويُقدر تقرير مؤشر الذكاء الاصطناعي أن انبعاثات ثاني أكسيد الكربون الناتجة عن تدريب نموذج واحد من نماذجه تعادل حوالي 100 ضعف الانبعاثات التي ينتجها الإنسان في عام كامل.

إن تطوير الذكاء الاصطناعي يضع ضغطاً لا يمكن تصوره على الموارد الطبيعية، من خلال استخدام كميات ضخمة من المياه لتبريد مراكز البيانات التي تغذيه[12]. كما أن تطويره يعوق التحول إلى استخدام الطاقة النظيفة؛ بسبب اعتماده الثقيل على المواد كثيفة الكربون مثل الخرسانة والألمنيوم والأتربة النادرة والليثيوم المستخدمين في صناعة السيارات الكهربائية. إن تقديرات جمعية الطاقة الدولية تشير إلى أن طاقة مراكز بيانات الذكاء الاصطناعي سترتفع إلى الضعف بحلول عام 2026، وهذا يعادل إضافة دولة صناعية جديدة، بثقل السويد أو ألمانيا إلى الكوكب.

تحيز ممارسات الذكاء الاصطناعي

بعد إلقاء الضوء على كيفية قيام الذكاء الاصطناعي بإعادة إنتاج عدم المساواة الاجتماعية، يوضح القسم أدناه بعض تحديات الحياة الواقعية المتعلقة بتطبيقات الذكاء الاصطناعي. ونظراً لأن الأمثلة أكثر مما يمكن إحصاؤه، سنكتفي بإعطاء بعض الأمثلة؛ مما من شأنه أن يرسم صورة لكيفية قيام الذكاء الاصطناعي بتوليد عدم المساواة والتمييز والتحيز.

> ”يُقدر تقرير مؤشر الذكاء الاصطناعي أن انبعاثات ثاني أكسيد الكربون الناتجة عن تدريب نموذج واحد من نماذجه تعادل حوالي 100 ضعف الانبعاثات التي ينتجها الإنسان في عام كامل.

على سبيل المثال، من بين 7 آلاف لغة يتم التحدث بها حول العالم، يقتصر تدريب النماذج الكبيرة للذكاء الاصطناعي على حوالي 100 لغة منها فقط. وتمثل اللغة الإنجليزية، التي يتحدث بها 16% فقط من سكان العالم، اللغة التي يبنى عليها تدريب نماذج الذكاء الاصطناعي، وثلثي اللغات المستخدمة في الإنترنت. وهكذا فإن اللغة الإنجليزية تقوم "بنقل قيمها وأفكارها إلى سياقات اللغات الأخرى حتى ولو أن هذه القيم والأفكار لا تتفق مع قيم وأفكار تلك اللغات"[13].

وكمثال آخر نجد أن هناك 300 لغة أصلية في منطقة الأمازون بالبرازيل، لم تعترف النماذج اللغوية الآلية (LLMs) إلا بعدد قليل فقط منها[14]. هناك وجه آخر للتحدي الذي تسببه النظرة العالمية "الأنجلومركزية" التي تفرضها النماذج اللغوية الآلية (LLMs) المدربة على اللغة الإنجليزية فقط، ألا وهو الفشل في التقاط التفاصيل والفروقات الدقيقة؛

والتي بدورها قد تؤدي إلى اتخاذ قرارات مصيرية مغيرة لحياة البعض، كرفض طلبات لجوئهم السياسي[15].

هناك وقائع كثيرة على كيفية قيام أنظمة الذكاء الاصطناعي بتكرار وتعزيز عملية صنع القرار العنصرية والمعادية للأقليات العرقية والدينية. فمثلاً حين تم استخدام هذه الأنظمة للتنبؤ بمعدلات إمكانية العودة للجريمة مرة أخرى في الولايات المتحدة، تنبأت خوارزميات الذكاء الاصطناعي بأن "المتهمين السود هم الأكثر عرضة لارتكاب جرائم عنف في المستقبل بنسبة تصل إلى 77%، وهم كذلك الأكثر عرضة لارتكاب جرائم في العموم بنسبة 45%"[16].

يُعد قرار الحصول على سكن مناسب مجالاً آخر لوحظت فيه عنصرية أنظمة الذكاء الاصطناعي؛ إذ يمكن لهذه الأنظمة أن "تحرم الأشخاص الملونين من فرص عادلة لإيجاد سكن"[17]؛ إذ وجد تحقيق في خوارزمية كيفية اتخاذ قرار الموافقة على الرهن العقاري أنه بالمقارنة مع المتقدمين من البيض، تم رفض المتقدمين من السود، والأمريكيين الأصليين، وسكان جزر المحيط الهادي/الآسيويين، واللاتينيين بمعدل 80% و70% و50% و40% على التوالي[18].

تشـير العديـد مـن الحـوادث الأخـرى إلى التمييـز الـذي يمارسـه الـذكاء الاصطناعـي ضـد النسـاء، والأقليـات العرقيـة والإثنيـة والدينيـة[19]، والمهاجريـن، والأشـخاص ذوي الإعاقـة، في الوصـول إلى خدمـات الرعايـة الصحيـة[20] العامـة[21]، والتوظيـف، والتمويـل[22]، وغيرهـا مـن الخدمـات بسـبب التحيـزات المتأصلـة التـي تتضمنهـا أنظمتـه المختلفـة.

نظـراً لأن الـذكاء الاصطناعـي أصبـح يطمـس بشـكل متزايـد الخطـوط الفاصلـة بـين العالـم الافـتراضي والعالـم الواقعـي (ممـا يغـير تجاربنـا الحياتيـة)، قامـت الحكومـات والـشركات بطـرح العديـد مـن السياسـات والمبـادرات والأطـر في محاولـةٍ منهـا لإدارة وإصـلاح مخاطـر الـذكاء الاصطناعـي، وذلـك بطرحهـا حلـولاً تقـوم بالاعتمـاد (المفـرط) عـلى التكنولوجيـا لمشـكلة قامـت في الأسـاس بسـبب التكنولوجيـا.

لقـد طغـى علينـا التوجـه نحـو المركزيـة التكنولوجيـة، ودفعنـا بالرؤسـاء التنفيذيـين وعلمـاء الـذكاء الاصطناعـي والمهندسـين إلى مسـتوى مرتفـع جـداً مـن التبجيـل، حتـى صـاروا بمثابـة قـادة اجتماعيـين وسياسـيين غـير منتخبـين، يفرضـون وجهـات نظرهـم حـول المسـتقبل عـلى البشريـة جمعـاء. إن هـؤلاء الأشـخاص ليـس لديهـم ترخيـص لتمثيلنـا؛ ومـن ثـم، فمـن الـضروري أن نعمـل جاهديـن لامتـلاك القـدرة عـلى التفويـض فيمـا يتعلـق بمسـتقبل الـذكاء الاصطناعـي، فمجـرد قدرتنـا عـلى بنـاء مثـل هـذه الأنظمـة، لا يعنـي بالـضرورة أننـا يجـب أن نقـوم ببنائهـا، خاصـة بمثـل هـذه التكاليـف الباهظـة التـي لا يمكـن العـدول عنهـا. وهنـا يقـع علينـا واجـب المعارضـة والرفـض لهـذه الأنظمـة المميـزة والعنصريـة وغـير العادلـة، واسـتبدالها بـرؤى بديلـة عادلـة ومنصفـة ومتاحـة وممثلـة للجميـع.

❋ ترجمة: آلاء البدري

الهوامش:

1- Safiya Umoja Noble, Algorithms of oppression: How search engines reinforce racism, (**New York University Press**, 2018).

2- Harini Suresh, John V. Guttag, A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle, **EAAMO'** 21, 2021. https://doi.org/10.1145/3465416.34833

3- Sarah Brayne, **Predict and surveil: Data, discretion, and the future of policing**, (Oxford University Press, 2020).

4- Sarah Brayne, S., & Angele Christin, Technologies of crime prediction: The reception of algorithms in policing and criminal courts. **Social Problems**, 68, 2021.

5- Mike Zajko, **Artificial intelligence, algorithms, and social inequality: Sociological contributions to contemporary debates**. Sociology Compass, 16(3), 2022. https://doi.org/10.1111/soc4.12962

6- Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, Aram Galstyan. **A survey on bias and fairness in machine learning**, 2022 ArXiv:1908.09635 [Cs]. http://arxiv.org/abs/1908.09635

7- Nestor Maslej, Loredana Fattorini, Erik Brynjolfsson, John Etchemendy, Katrina Ligett, Terah Lyons, James Manyika, Helen Ngo, Juan Carlos Niebles, Vanessa Parli, Yoav Shoham, Russell Wald, Jack Clark, and Raymond Perrault, "**The AI Index 2023 Annual Report**," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2023. https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf

8- Ben Cottier, Robi Rahman, Loredana Fattorini, Nestor Maslej, and David Owen. '**The rising costs of training frontier AI models'**. ArXiv [cs.CY], 2024. arXiv. https://arxiv.org/abs/2405.21015

9- **Ibid**.

10- Nestor Maslej et. al., "The AI Index 2023 Annual Report," **AI Index Steering Committee, Institute for Human-Centered AI,** Stanford University, Stanford, CA, April 2023. https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf p.171.

11- Somya Joshi, Björn Nykvist, **Anticipating Futures: How Artificial Intelligence acts as an Amplifier of Inequity**, Stockholm Environment Institute, 2023. https://tinyurl.com/dczrcxcd

12- Financial Times, **AI boom sparks concern over Big Tech's water consumption**, Cristina Criddle & Kenza Bryan, 25 February 2024. https://tinyurl.com/8uknsecv

13- Gabriel Nicholas and Aliya Bhatia, Lost in Translation: **Large Language Models in Non-English Content Analysis**. Center for Democracy & Technology, 2023. https://tinyurl.com/2zderytm

14- Madeleine North, **Generative AI is trained on just a few of the world's 7,000 languages. Here's why that's a problem – and what's being done about it**, World Economic Forum, 2024. https://www.weforum.org/agenda/2024/05/generative-ai-languages-llm/

15-Carey L. Biron, **AI's 'insane' translation mistakes endanger US asylum cases,** Context News, 2023. https://tinyurl.com/2s4jjtzw.

16- Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, **Machine Bias There's software used across the country to predict future criminals. And it's biased against blacks,** ProPublica, 2016, https://tinyurl.com/2tnzs9ce

17- Khari Johnson, **Algorithms Allegedly Penalized Black Renters. The US Government Is Watching, Wired, 2023**. https://tinyurl.com/8fy3hjny

18- Emmanuel Martinez and Lauren Kirchner, **The Secret Bias Hidden in Mortgage-Approval Algorithms**, The Markup, 2021, https://tinyurl.com/52wa6htr

19- Nenad Tomasev, Jonathan Leader Maynard, & Iason Gabriel, **Manifestations of xenophobia in AI systems**. AI & Soc (2024). https://doi.org/10.1007/s00146-024-01893-4

20- Shazeen Suleman, Kent D. Garber & Lainie Rutkow, **Xenophobia as a determinant of health: an integrative review**. J Public Health Pol 39, 407–423 (2018). https://doi.org/10.1057/s41271-018-0140-1

21- Amnesty International, **Xenophobic Machines Discrimination Through Unregulated Use Of Algorithms In The Dutch Childcare Benefits Scandal**, 2021. https://www.amnesty.org/en/latest/news/2021/10/xenophobic-machines-dutch-child-benefit-scandal/

22- Ryan Browne and MacKenzie Sigalos, **A.I. has a discrimination problem. In banking, the consequences can be severe**, CNBC, 2023.https://tinyurl.com/2s3689ra