



# تجريم التحيز الخوارزمي وآليات الحد منه (دراسة مقارنة)

**الدكتور**

**أحمد كامل نعيم**

**دكتوراه في القانون الجنائي**

**ahmedmoneim106@gmail.com**

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

## المخلص:

تعتبر ظاهرة التحيز الخوارزمي من أبرز التحديات القانونية والأخلاقية التي تواجه المجتمع المعاصر، خاصة مع التطور المتسارع للذكاء الاصطناعي وتطبيقاته الواسعة في شتى المجالات حيث تصاعدت حدة التحيز في الخوارزميات تزامناً مع انتشار استخدامها من الهيئات الحكومية والخاصة والشركات والمنصات الإلكترونية ومحركات البحث الذكية والشركات التكنولوجية الكبرى بمناسبة عملها في تحليل وفلترة البيانات الضخمة على الإنترنت في مجالات عديدة لاسيما مجال العمل والتوظيف والرعاية الصحية، والائتمان، والعدالة الجنائية التنبؤية، والتسويق الإلكتروني.

مما يجعل بلا شك وجود تجريم للتحيز الخوارزمي سواء كان التحيز من أفعال المعلومات أو تحيز في النتائج التي تظهرها تلك الحواسيب؛ ويتسع نطاق مخاطر التحيز الخوارزمي ليس فقط لسرعة انتشار استخدام الخوارزميات وإنما لانعدام الشفافية حول آلية عملها وسرية تصميمها وصعوبة رقابتها أو تدقيقها باعتبارها مملوكة لشركات خاصة وتحفظ بملكيته وأدارتها وتغذيتها.

ولما كان القانون هو انعكاس للظواهر المجتمعية فتنامي ظاهرة التحيز الخوارزمي تدعو إلى التساؤل حول مدى صلاحية القواعد القانونية الجنائية التقليدية للتطبيق عليها أم أن هذه الصورة الحديثة من صور التمييز العنصري تتطلب تدخلا تشريعيا لإضفاء صفة التجريم عليها بغية إعلاء مبدأ المساواة ليس فقط في الواقع المادي وإنما في بيئة الإنترنت أيضاً.

**الكلمات المفتاحية: التحيز الخوارزمي العدالة الخوارزمية، التمييز العنصري، قانون مسألة**

**الخوارزميات، قانون الخدمات الرقمية الأوروبي.**

## **ABSTRACT**

The phenomenon of algorithmic bias is one of the most prominent legal and ethical challenges facing contemporary society, especially with the rapid development of artificial intelligence and its wide applications in various fields, as the intensity of bias in algorithms has escalated in conjunction with the spread of their use from government and private bodies, companies, electronic platforms, smart search engines, and major technological companies as they work to analyze and filter huge data on the internet in many fields, especially the field of labor, employment, health care, credit, predictive criminal justice, and e-marketing. The risk of algorithmic bias increases not only because of the rapid spread of the use of algorithms, but also because of the lack of transparency about how they work, the secrecy of their design, and the difficulty of monitoring or auditing them as they are owned, managed, and fed by private companies. Since law is a reflection of societal phenomena, the growing phenomenon of algorithmic bias calls into question the validity of traditional criminal legal rules to apply to it, or whether this modern form of racial discrimination requires legislative intervention to criminalize it in order to uphold the principle of equality not only in the physical reality but also in the line environment

**Keywords: Algorithmic Bias, Algorithmic Justice, Racial Discrimination, Racial Discrimination, Algorithmic Law, European Digital Services Act**

## مقدمة:

لئن كانت الثورات الصناعية السابقة قد أتاحت للإنسان استبدال عضلاته في العمل بالأدوات البسيطة والمتطورة، فإن الثورة الصناعية الرابعة قد مكنته من استخدام ذكاء اصطناعي عوضاً عن ذكائه البشري في إدارة شتى شؤون حياته الشخصية، والعلمية، والطبية، والتجارية... إلخ، حيث استطاع هذا الذكاء الخروج من صفحات روايات الخيال ومن مختبرات البحوث إلى الواقع المادي الملموس بفضل الخوارزميات والبرمجيات والتقنيات المتطورة التي تستعيز عن ذكاء الإنسان البشري إلى ذكاء مصطنع يحل مكانه في كافة نواحي الحياة القانونية.

الحقيقة إن معظم الدول قد أدركت عظم التحديات التي تمثلها تقنيات الذكاء الاصطناعي مستقبلاً وكذلك في القرارات التي قد تتم اتخاذها عبر تلك الأدوات؛ فسارعت إلى إنشاء الوزارات والمجالس القومية والكلية التي تعمل في هذا الميدان، حيث أحدثت الثورة التكنولوجية اختلافاً جوهرياً في المجالات الحياتية المختلفة سواء على صعيد الفرد أو الجماعة.

لا يخفى على أحد اليوم أن الذكاء الاصطناعي قد أدى إلى قلب الموازين على كل الأصعدة، حيث لجأت أغلب الدول إلى اعتماده في شتى قطاعاتها العسكري، والطبي، والتعليمي والتوظيفية... إلخ، بل وحتى الأفراد في حياتهم العلمية والعملية، وهذا نظراً لقدرته على تنفيذ وحل كل المهام والمسائل ذات التعقيدات بالنسبة للبشر وخصوصاً في بعض الآليات التي يستخدمها أدوات الذكاء الاصطناعي التي قد تؤدي إلى مشكلة التحيز الخوارزمي في كافة أنحاء والمجالات المستخدمة في الذكاء الاصطناعي.

وعلى الرغم من كل ما قدمته أنظمة الذكاء الاصطناعي من منافع للبشرية، إلا أنها في الجهة المقابلة صدمتها بوجود تغذية تحيزه في خوارزمياتها<sup>(1)</sup>، أو ما يعرف بالتحيز الخوارزمي في مجال العدالة الجنائية وأثرها على الحقوق والحريات العامة، الذي يقوم بالانحياز لطائفة على حساب طائفة أخرى دون الاستناد إلى مبررات موضوعية أو منطقية بل يقوم بتحيز بناء على ادخال معلومات ذات صبغة تحيزه من قبل مبرمجها<sup>(2)</sup>.

---

(1) Songül Tolan, Fair and Unbiased Algorithmic Decision Making: Current State and Future Challenges, European Commission, Joint Research Centre, December 2018, Accessed 20-8-2023, On this website;

<https://arxiv.org/ftp/arxiv/papers/1901/1901.04730.pdf>

(2) يمكن تعريف الذكاء الاصطناعي (A.I) - Artificial Intelligence - ما هو إلا محاكاة للذكاء البشري، عن طريق القيام بالعمل على تطوير تقنيات تكنولوجية، تكون قادرة على أداء العطاءات البشرية بشكل ذكي عن طريق استخدام احتمالات المنطق من خلال البيانات.

وهكذا، فإن لجوء القطاعات العامة أو الخاصة على السواء إلى الاعتماد على أنظمة الذكاء الاصطناعي المغذية بالتحيز في القيام بالوظائف واتخاذ القرارات وفي مجال العدالة الجنائية عامة، فيه مساس بحقوق الإنسان على وجه الخصوص، منها حق العمل والحياة والتوظيف على وجه الخصوص، لكونه الحق الذي يضمن له العيش بكرامة، في الوقت الذي كنا ننتظر فيه أن يقوم الذكاء الاصطناعي بتعزيز حقوق الإنسان وحياته، وذلك بإزالة الفوارق العرقية والجنسية والثقافية بين الأفراد والشعوب.

حيث يمكن تعريف العدالة الخوارزمية بأنها الاستخدام بطريقة منصفة وعادلة للخوارزميات المستخدمة في العدالة الجنائية والحرية الأساسية التي تمس الإنسان وحيته بصفه اساسية، وأضحت تُمثل تحدياً في العصر الذي نعيشه في العصر الرقمي التي طغت عليه استخدام خوارزميات الذكاء الاصطناعي<sup>(1)</sup>. فمع تزايد استخدام تلك الخوارزميات وأنظمة الذكاء الاصطناعي وأشارته في عمليات اتخاذ القرارات الحيوية والمصيرية، حيث أصبح من اللازم ضرووه وضع اليات ضمان عدم وجود تحيز أو تمييز غير عادل بين الأشخاص سواء في المدخلات أو في النتائج التي تصدرها تلك الأنظمة بُغية تحقيق مستقبل تقني أكثر عدلاً وتكافؤاً.

غير أن الواقع العملي الذي كشف العديد من النماذج الواقعية وجود التحيز الخوارزمي الذي أثر بشكل مباشر وواضح في الحياة الواقعية ومن ضمن تلك المجالات التمييز والتحيز العنصري بواسطة خوارزميات الذكاء الاصطناعي أو اليات العنصر الألى، الأمر الذي بات يهدد مبدأ المساواة بين الأشخاص ويجعل العدالة الخوارزمية محل تأثيم جنائي حيث تعددت وسائل تطويع خوارزميات الذكاء الاصطناعي في مجالات شات، تبيّن لاحقاً مدى تحيزها لعرق معين أو دين بعينه أو جنس دون الآخر، ولعلّ أبرز هذه النماذج، خوارزميات بعض نُظم لاستدلالات الجنائية المتحيّزة تجاه بعض الفئات في المجتمع والتي تعمل من خلالها الشرطة التنبؤية للقياس مد خطورة المتهم مما أدّى إل اعتقال غير منصف أو توجيه اتهام غير قانونيا لبعض الأشخاص بناءً على سماتهم الشخصية، واهم تلك الأنظمة في التطبيق برنامج "كومباك COMPAS الذي استخدم خوارزميات متحيّزة<sup>(2)</sup> ضد بعض الأمريكيين

---

Brigitte Lasry & Hael Kobayashi, Human Decisions Thoughts On AI, The united nations educational, Scientific and cultural organization, NETEXPLO, 2018, pp. 20-21.

(1) Chouldechova, Fair prediction with disparate impact: a study of bias in recidivism prediction instruments. February 8, 2017, Accessed 20-8-2023, On this website; [https://www.andrew.cmu.edu/user/achoulde/files/disparate\\_impact.pdf](https://www.andrew.cmu.edu/user/achoulde/files/disparate_impact.pdf)

(2) حكمت محكمة العليا في ولاية وسكونسن الأمريكية على متهم في عام 2016 بالسجن لمدة ست سنوات مؤسسة حكمها على نتائج تقرير التطبيق الخوارزمي، "COMPAS" الاطلاع على الحكم على الرابط التالي؛ بتاريخ دخول

٥-٩-٢٠٢٣:

ذوى البشرة السواء أو من أصل أفريقي أو آسيوية، وصافنتهم في درجة كبير من الخطورة الإجرامية مقارنة بالأمريكيين البيض الأمر الذي أدى إل بقائهم مدد حبس احتياطي أطول بدون ميرر قانونيا ورفض الأفراج المشروط بناء على معلومات متحيزه تم أدخلها للنظام المستخدم أدى الى نائج متحيزة (١).

من الجدير بالذكر، أن الخوارزميات الذكاء الاصطناعي ودورها في مجال العدالة التنبؤية ارتبطت بشكل كبير بكيفية استخدام وتوظيف البيانات القضائية الضخمة، واستخدام تلك المعطيات في الدراسة والتحليل بواسطة خوارزميات قد تتميز بالتحيز الخوارزمي، واعتمادها على برمجيات وبرامج صممت خصيصاً لهذا الغرض وخاصة ان استعمال تلك التقنيات ستؤثر حتماً على طريقة صنع القرارات والأحكام القضائية واختيار المقدمين لشغل الوظائف العامة والتميز بين الجنسين بدون مبرر يذكر، حيث أن من اعظم دواعي استخدام تلك الخوارزميات عدم وجود تميز وتحيزوا لكن واقع الأمر غير ذلك مع الاستخدام المفرط لتلك الخوارزميات؛ إذ أصبحت العديد من الدول تعتمد على تقنية الخوارزميات في تقديم بعض من خدماتها الخاصة أبرزها بقطاع العدالة الجنائية، فعلى سبيل المثال، تستخدم في إصدار الأحكام القضائية على أساس السوابق القضائية للمتهمين، وتحديد ضرورة وضع مشتبه به بالحبس الاحتياطي من عدمه؛ انطلاقاً من تحليل البيانات المتعلقة باستعمال الخوارزميات... إلخ.

أظهرت كذلك دراسة نشرتها مجلة "splinter" أجريت على ٢٠٠ مليون شخص في الولايات المتحدة الأمريكية في مجال الرعاية الصحية والتوظيف حيث أظهرت الت تحيزاً عنصرياً ضد ذوى البشرة السوداء، حيث صنفتهم الخوارزميات في درجة أقل من حيث المخاطر الصحية وفي مجال التوظيف على غير الحقيقة وهو ما جعلهم عرضة لحجب الرعاية الصحية وأقل بأقرانهم في التسلسل الوظيفي عنهم فضلاً عن تلقيهم رعاية أقل جودة مقارنة بذوى البشرة البيضاء. (٢)

لم تلك الخوارزميات المتحيزة قاصرة على التمييز العنصري بين الأشخاص على أساس العرق فقط، بل امتدت كذلك للتمييز بينهم على أساس السن، ففي فبراير ٢٠٢٣ رفعت دعوى قضائية جماعية

---

<https://law.justia.com/cases/wisconsin/supreme-court/2016/2015ap000157-cr.html>

(1) Corbett-Davies, Sam, Emma Pierson, Avi Feller, Sharad Goel, and Aziz Huq. "Algorithmic Decision Making and the Cost of Fairness." ArXiv:1701.08230 [Cs, Stat], January 27, 2017. Accessed 20-8-2023, On this website;

<https://arxiv.org/abs/1701.08230>.

(2) Starre Vartan; Racial Bias Found in a Major Health Care Risk Algorithm, Black patients lose out on critical care when systems equate health needs with costs, splinter, 2019, Accessed 20-8-2023, On this website;

[https://link.springer.com/chapter/10.1007/978-3-031-46138-5\\_10](https://link.springer.com/chapter/10.1007/978-3-031-46138-5_10).

لدى محكمة كاليفورنيا بالولايات المتحدة الأمريكية ضد شركة (Workday) التي تضطلع بتقييم طلبات التوظيف لأكثر من ٥٥ مليون مستخدم، وتعتمد عليها أكثر من ٥٠٠ شركة في التوظيف<sup>(١)</sup>، حيث رفضت خوارزمياها تعن الأشخاص الذين يزيد أعمارهم على ٤٠ عاماً، كذلك المتقدمين من ذوى الهمم القانون الحقوق المدنية الأمريكي لعام ١٩٦٤ الذى يحظر على وجه الخصوص التمييز في التوظيف<sup>(٢)</sup>، وقانون الأمريكيين ذوي الإعاقة المعدل في عام ٢٠٠٨<sup>(٣)</sup>

قد أدى التحيز في الخوارزميات الى التأثير على كثير من الأنشطة البنكية كذلك الوضع الائتماني للشركات وللأفراد، حيث واجهت الشركات المتعثرة والأشخاص من ذوى الدخل المنخفض والأقليات في الولايات المتحدة الأمريكية تحيزاً في عمليات الثقة المصرفية ونظام النقاط البنكية التي تؤثر بشكل رئيسي على منح القروض المصرفية حيث تمت التقييمات بناء على خوارزميات تقيم الوضع المالي المبنية على عوامل معينة تؤدي إلى تمييز على أساس العرق أو الجنس أو العنوان البريدي وسجلهم المالي<sup>(٤)</sup>، ما يؤثر على فرص حصولهم على القروض المتنوعة بشكل غير عادل وحظرهم من الوصول إلى سوق الرهن العقاري وهو ما يخل بمبدأ تكافؤ الفرص الائتمانية.

وفي الشركات الكبرى والأسواق الإلكترونية وجد التحيز الخوارزمي موطأ قدم له حيث كانت خوارزميات شركة Apple متحيزة جنسياً ضد النساء فيما يتعلق بالحدود الائتمانية التي تمنحها إدارة الخدمات المالية (Apple Card) والتي انخفضت ما بين ١٠ أضعاف إلى ٢٠ ضعف بالنسبة للنساء

---

القضية نيابة عن نفسه وعن آخرين في وضع مماثل، بما في ذلك جميع المتقدمين الأمريكيين Mobley ارفع موبلي<sup>(١)</sup> عاماً، والمتقدمين ذوي 40 من أصل أفريقي، بما في ذلك المتقدمون السابقون، والمتقدمون الذين تزيد أعمارهم عن للتوظيف نتيجة 2023. وفبراير 2019 يونيو 3 الإعاقة الذين لم تتم إحالتهم أو تعيينهم بشكل دائم في الفترة ما بين ضد شركة Mobley لعملية الفحص التمييزي. الاطلاع على مذكرة الدعوى المقدمة من ديريك موبي وشركات أخرى، المزمع الحكم فيها في مطلع عام، ٢٠٢٤ على الرابط التالي؛ بتاريخ دخول: ٢٠٢٣-٩-Workday-  
<https://www.courtlistener.com/docket/66831340/1/mobley-v-workday-inc/>

<sup>(٢)</sup> متاح للأطلاع ديسمبر ٢٠٢٤ على الرابط التالي - تاريخ الاطلاع ١/١٢/٢٠٢٤ .

<https://www.eeoc.gov/statutes/title-vii-civil-rights-act-1964>

<sup>(3)</sup> [discrimination/EFFAB30A520782135D1EFF291F32E22E/card-](https://www.cambridge.org/core/journals/business-discrimination/EFFAB30A520782135D1EFF291F32E22E/card-)

<https://www.cambridge.org/core/journals/business> <https://www.apple.com/apple>

<sup>(4)</sup> Garcia, A.C.B., Garcia, M.G.P. & Rigobon, R. Algorithmic discrimination in the credit domain: what do we know about it?. *AI & Soc* (2023). Accessed 20-8-2023, On this website; <https://link.springer.com/article/10.1007/s00146-023-01676-660>. -Elegido, J. The Ethics of Price Discrimination. *Business Ethics Quarterly*, Vol.21, I.4, 2015, pp.633 (4) Accessed 20-8-2023, On this website;-price-of-quarterly/article/ethics-ethics-

عن نظيرتها الممنوحة للرجال رغم تماثل ووحدة المركز المالي لكل منهم، ما يعني انحياز الخوارزميات ضد المرأة<sup>(1)</sup>.

هذه الأمثلة ما هي إلا غيض من فيض التحيز الخوارزمي المستشري والذي يزداد استخدامه يوماً بعد يوم في مجالات عديدة حتى في القضاء الجنائي، ليجعل التمييز غير العادل بين الأشخاص أمراً ممنهجاً، وتجريمه صراحة اضحى أولوية نظراً للمصلحة الجوهرية التي يحميها، والتي تؤثر على القيمة القانونية لمبدأ المساواة الذي لم يعد قاصراً في مفهومه التقليدي على عدم التمييز المادي بين الأفراد، بل يشمل أيضاً عدم التمييز في بيئة الذكاء الاصطناعي أملاً في تحقيق العدالة

من خلال تسليط الضوء على بعض أوجه القصور في اتخاذ القرار الخوارزمي، وتحديد المواضيع الرئيسية التي تدور في فلك مشكلة الأخطاء والتحييزات الخوارزمية مثل تغذية البيانات والتأثير المتباين للخوارزميات، وتناول بعض المقاربات لمواجهة هذه الإشكاليات، وتعد هذه الدراسة على قدر كبير من الأهمية لصناع القرار والمختصين الذين يسعون إلى اكتساب فهم أفضل لكيفية تأثير تحيز خوارزميات الذكاء الاصطناعي على أصحاب الشأن لديهم، وهذا من شأنه التأثير على مؤسسات العدالة الجنائية، وبالنظر إلى الطبيعة الخاصة لنظام الخوارزميات وتصميمها الرقمي حيث تتميز بالغموض والتعقيد التنبؤي مما يجعل من الصعوبة بمكان تقييم الحكم على ادائها، وتقدير مخاطرها كما يمكن لغموضها أن يحجب تسبب القرارات والأحكام المؤتمتة، لذا تعد هذه الدراسة مدخلاً للمختصين بما تتناوله من أهم العناصر الجوهرية لمفهوم "خوارزميات الذكاء الاصطناعي" والذي يقوم على الربط بين جانب قانوني وجانب تقني علمي.

إشكالية الدراسة:

إذا كانت تقنيات الذكاء الاصطناعي وتطبيقاتها، وما يرتبط بكل الطفرات التكنولوجية الجديدة، أصبح يمثل تحدياً حتى على المستوى القانوني والأخلاقي<sup>(2)</sup>، فإنه بالنظر إلى أهمية خوارزميات الذكاء الاصطناعي، والطبيعة الخاصة التي تتميز بها، فإن الإشكالية الأساسية التي تطرحها الدراسة: كيف

---

(1) Following complaints of gender discrimination, Apple Card is being investigated. On Twitter, a well-known software engineer claimed that the credit card was "sexist" since it discriminated against women who applied for credit. On this website: <https://www.nytimes.com/2019/11/10/business/Apple-credit-card-investigation.html>, accessed June 9, 2023

(2) De Souza, M. A. L., Intelligence artificielle et philosophie: les critiques de H. L. Dreyfus et J. Searle à intelligence, Mémoire, Philosophie, Université du Québec à Trois-Rivières, mars, 1992, p.61.

يمكن لخوارزميات الذكاء الاصطناعي المغذية بالتحيز إصدار قرارات تتصف بالتمييز المجرم جنائياً للتمييز بين الأفراد في مجال العمل والتوظيف وحقوق الإنسان والعدالة الجنائية؟

وتتفرع عن الإشكالية الرئيسية مجموعة من التساؤلات الفرعية يمكن إجمالها في النقاط الآتية: ما المقصود بالتحيز الخوارزمي ما هي طبيعته القانونية؟ وماهي الأسباب المؤدية إلى التحيز الخوارزمي هل التحيز من فعل الآلة أو من تعليم الإنسان للآلة؟ ماهي أنواع التحيز الخوارزمي؟ هل يمكن للتحيز الخوارزمي التأثير على القرارات والاحكام المستخدمة فيها الذكاء الاصطناعي بما ينطوي على التمييز بدون مسوغ قانوني أو أخلاقي وآليات الحد من التحيز الخوارزمي؟

كيف تناولت التشريعات الوطنية والمقارنة ظاهرة التحديات القانونية والأخلاقية للتحيز الخوارزمي التي بدت في الانتشار حديثاً؟ وما هو موقف المشرع المصري من تجريم تلك الظاهرة وهل تستغرق نصوص قانون العقوبات التقليدية هذا السلوك أم يظل مشروعاً لحين تدخل المشرع بتجريمه صراحة؟ ما المقصود بمشروعات قوانين المساءلة الخوارزمية التي نهضت في الولايات المتحدة في السبع سنوات الأخيرة؟ وكيف تناولت هذه المشروعات آليات الحد من ظاهرة التحيز الخوارزمي؟ وهل جرمتها أم اكتفت بالحظر؟

أهمية الدراسة:

تأتي أهمية هذا الموضوع بالأساس من الواقع العملي الذي يبرز فيه التحيز الخوارزمي المستمد من لاستخدامات الذكاء الاصطناعي للبرامج والمواقع التوظيف والتمييز بين الجنسين، وما أحدثته من انتهاكات صارخة لحقوق الافراد مما كشفه الواقع من تراجع في الحياد الخوارزمي بسبب الغموض الذي يكتنف تصميم الخوارزميات وكيفية عملها من فرز وتحليل وتصنيف واتخاذ قرارات آلية قد تكون متحيزة لعرق معين أو سن محدد أو غيرها في نهاية المطاف، باعتبار أن سريتها حقاً تجارياً استثنائياً على صانعيها. هذه الضبابية وانعدام الشفافية تجعل إلقاء الضوء على موضع الخوارزميات المتحيزة في القانون الجنائي أمراً مهماً للغاية لتعلقها بمضمون القانون وهو العدالة استشهادا بمقولة الفيلسوف البريطاني إدموند بيرك-Edmund Burke "حيث يبدأ الغموض تنتهي العدالة"<sup>(1)</sup>. من ثم يكون للبحث أهمية موضوعية تتعلق لكيفية تجريم التحيز الخوارزمي من الناحية القانونية والأخلاقية وأثره في المجال الرقمي من خلال اسقاط نصوص قانون العقوبات المصري والأمريكي على هذا السلوك غير المشروع

(1) A good parson once said that where mystery begins religion ends. Cannot I say, as truly at least, of human Edmund Burke (1729-97), Irish philosopher, statesman. A laws, that where mystery begins justice ends? Society Natural of website; 12Vindication this on -9-2023,

Accessed (1756). <http://www1.udel.edu/CRJU/dgulick/quote.htm>

فضلاً عن تناول الاتفاقيات الدولية التي بإمكان نصوصها استيعاب هذا النمط المستحدث من التمييز العنصري. كذلك للبحث أهمية إجرائية تتبدى من خلال تناول التشريعات الآليات المطلوبة لإلزام المنصات الإلكترونية على الإنترنت ومحركات البحث التي تتعامل مع البيانات الضخمة بالشفافية والتصحيح المطلوب من ظاهرة التحيز الخوارزمي.

### نطاق البحث:

يتحدد نطاق البحث حول تجريم التحيز الخوارزمي واثره في المجال الرقمي باعتباره صوره من صور التمييز العنصري بين الأشخاص، ومدى صلاحية المواثيق الدولية كمصدر لتجريم هذه الظاهرة، فضلاً عن بيان مسلك الدول الأوروبية في معالجة التحيز الخوارزمي وموقف مصر والولايات المتحدة الأمريكية منها، وبالتالي يخرج عن نطاق البحث كل ما يتعلق تعريف الذكاء الاصطناعي وأهميته واتجاهاته إلا إذا كانت في معرض الحاجة، بيان كذلك يخرج عن البحث الصور الأخرى للتمييز العنصري كالحض على الكراهية والعنف وازدراء الأديان وغيرها من الصور.

### منهجية البحث:

في سبيل تبسيط موضوع البحث والإجابة عن التساؤلات المطروحة، فإن الباحث سوف يقسم موضوع الدراسة الى قسمين: أولهما: هو المنهج الوصفي، في تعريف ماهية التحيز الخوارزمي، وأسبابه وأنواعه، وثانيهما: هو المنهج الاستقرائي وذلك عبر استقراء كيفية استنباط اثار التحيز الخوارزمي واثره على التمييز المجرم جنائياً، وآليات الحد منه، إلى جانب استقراء توصيات المنظمات بشأن أخلاقيات الذكاء الاصطناعي وقانون الخدمات الرقمية الأوروبي لعام ٢٠٢٢ من تجريم ظاهرة التحيز الخوارزمي الذي تعد تحدى قانوني وأخلاقي لاستخدام تلك الخوارزميات.

### خطة الدراسة:

تم تقسيم الدراسة إلى مبحثين، تناول في المبحث الأول مفهوم التحيز الخوارزمي في أنظمة الذكاء الاصطناعي، وأسبابه وأنواعه، أما المبحث الثاني، فقد جاء تحت مسمى أثر تغذية خوارزميات الذكاء الاصطناعي بالتحيز على الحق في التوظيف والعدالة الجنائية، والذي عالج تأثير التحيز الخوارزمي على الحق في العمل، وآليات الحد منه والذي سوف نتعرض لها كما يلي.

## المبحث الأول ماهية التحيز الخوارزمي

تمهيد وتقسيم:

يمكن تعريف مصطلح الخوارزمية: أنه "عملية أو مجموعة من العمليات التي يجب اتباعها في حل المشكلات، فهي عملية منظمة، تتابع في خطوات منطقية"، وأيضًا يعرف بأنه: "مجموعة من العمليات الرياضية والمنطقية والمتسلسلة اللازمة لحل مشكلة ما"<sup>(1)</sup> كذلك يمكن تعريفها<sup>(2)</sup> بأنها كل: "سلوك وخصائص معينة تتسم بها البرامج الحاسوبية تجعلها تحاكي القدرات الذهنية البشرية وأنماط عملها"، أي إنه قدرة الآلات الرقمية وأجهزة الكمبيوتر على التفكير والتعلم والاستنتاج واتخاذ القرار، ورد الفعل، واستخدام الخبرة المكتسبة، واتخاذ القرارات للقيام بمهام تحاكي تلك التي يقوم بها الإنسان بوعي وذكاء<sup>(3)</sup>. كما يمكن أن يعبر عنه بأنه نظام برمجيات قادر على تقليد طرق التفكير البشرية، ومحاكاة السلوك الإنساني بمساعدة الأجهزة الحاسوبية بواسطة استخدام الوسائل الرياضية والمنطقية. وباختصار شديد

---

(1) Manheim, K. & Kaplan, L., **Artificial Intelligence: Risks to Privacy and 2019, Democracy**, p.113,

[https://papers.ssrn.com/so13/papers.cfm?abstract\\_id=3273016](https://papers.ssrn.com/so13/papers.cfm?abstract_id=3273016); Wickersham, G., & al. **IBA Global Employment Institute Artificial Intelligence and Robotics and their Impact on the Workplace**, 2017, p.10.

وراجع حول هذه التعريفات سيد طنطاوي محمد الجوانب القانونية للذكاء الاصطناعي والروبوت، بحث منشور على موقع دورية المركز الديمقراطي العربي، ٢٩ فبراير ٢٠٢٠، على الرابط الإلكتروني التالي:

<https://democraticac.de/?p=64965>

Robotik, B., (2016). **Intelligent Agents and Criminal Law – Negligence, Diffusion of Liability and Electronic Personhood, Robotics and Autonomous Systems** 86; Russell S. J., & Norvig, P., **Artificial Intelligence: A Modern Approach**, Pearson Education Limited, 3rd edition, 2014, pp.14–17; Pham, Q. V., Nguyen, D. C., Hwang, W. J., & Pathirana, P. N. (2020). **Artificial Intelligence and Big Data for Coronavirus (COVID–19) Pandemic: A Survey on the State-of-the-Arts**, p.5; Bbalan, S. M., **Artificial Intelligence and Law: A Review of the Role**

(2) Heudin, Jean– Cl., **Les créatures artificielles: des automates aux mondes virtuels**, éd., O. Jacob, Paris, 2007, p. 75.

د. عادل عبد النور بن عبد النور، مدخل إلى عالم الذكاء الاصطناعي، المرجع السابق، ص ٧ وما بعدها.

نحن أمام محاولة إكساب برنامج صفات بشرية أهمها الذكاء، وما يرتبط به من تفكير وإبداع وتخطب وتعلم حاسوبي ذاتي<sup>(١)</sup>.

حيث يمكن تقسيم هذا المبحث الى مطالبين رئيسين ونقسمه الى المطلب الأول: تعريف ماهية التحيز الخوارزمي وأنواعه وطبيعته وأركانه القانونية وفي المطلب الثاني: أثر التحيز الخوارزمي على العدالة وحق التوظيف.

## المطلب الأول

### ماهية التحيز الخوارزمي "Algorithmic bias"

يُقسم مصطلح "التحيز الخوارزمي" لغوياً إلى كلمتين، الأولى هي "التحيز"، ومصدرها تحيز، وفعلها تحيّر إلى، فهو مُتَحَيَّرٌ، ويقال تحيّر إلى فلان أي حاباه وانضم إليه و وافقه في الرأي دون مراعاة للحق، والتحيز أذن هو الميل والمحاباة والتفضيل والتمييز، ويقال تحيّر الحُكْمُ لفریقٍ بلائِهِ: أي حاباهُ، ساندَهُ، تَشَيَّعَ لَهُ<sup>(٢)</sup>، وفي التنزيل العزيز: "وَمَنْ يُؤْمِدْ يُؤْمِدْ دُبْرَهُ إِلَّا مُتَحَرِّفًا لِقِتَالٍ أَوْ مُتَحَيِّرًا إِلَىٰ فِتْنَةٍ فَقَدْ بَاءَ بِغَضَبٍ مِنَ اللَّهِ"<sup>(٣)</sup> وفي تفسير هذه الآية الكريمة يقصد بالمتحيز إلى فئة هو الذي يفر إلى المسلمين ويعاونهم<sup>(٤)</sup>.

وهكذا، فإن هذا المطلب سوف يقسمه الباحث الى عدة فروع رئيسية من خلالها سنتعرض الى تعريف ماهية التحيز الخوارزمي المستخدم في أنظمة الذكاء الاصطناعي وأسبابه وطبيعته القانونية في الفروع الآتية:

(١) الخوارزمية: "مجموعة من التعليمات المحددة لحل مشكلة، أو أداء مهمة معينة" معجم البيانات والذكاء الاصطناعي إنجليزي-عربي، ط1، الهيئة السعودية للبيانات والذكاء الاصطناعي، مجمع الملك سلمان العالمي للغة العربية، 2022م، ص40.

(٢) معجم المعاني الجامع الإلكتروني، على الرابط التالي، بتاريخ دخول ٤-٩-٢٠٢٣:

<https://www.almaany.com/ar/dict/ar-ar/%D8%AA%D8%AD%D9%8A%D8%B2/>

(٣) سورة الأنفال - الآية ١٦.

(٤) ابن كثير: تفسير القرآن العظيم، خرج أحاديثه هاني الحاج، مراجعة ناصر الدين الألباني، المجلد الثاني، الجزء الرابع، المكتبة التوفيقية، طبعة ٢٠١٧، ص ١٦. (٤) كان الباحث الإسلامي في العصور الوسطى أبو جعفر محمد بن موسى الخوارزمي، الذي نسبت إليه الخوارزمية، أكثر اهتماماً بإجراءات محاسبية تتم خطوة بخطوة موثوقة لحلول الحوسبة للمعادلات الجبرية، وعلى مر العصور تحول مفهوم الخطوات الموثقة إلى الوظائف الحسابية **Computability and Computable Functions** لإضفاء الطابع المعتمد لفكرة الخوارزمية **Algorithm** وأخيراً وصل المفهوم إلى سلسلة محدودة من الإرشادات الدقيقة القابلة للتطبيق على أنظمة الكمبيوتر **Computing systems**.

## الفرع الاول

### تعريف التحيز الخوارزمي

#### في أنظمة الذكاء الاصطناعي

يقصد بتعريف التحيز الخوارزمي<sup>(1)</sup>: "انحياز أو تفضيل فئة على فئة في أنظمة الذكاء الاصطناعي مجموعات على أخرى بدون مبرر"<sup>(2)</sup>. وعرف أيضا أنه: "حالة انحياز وانحراف في النتائج المستخدمة خوارزميات التعلم الآلي يحدث بسبب وجود فرضيات متحيزة أثناء عملية تطوير الخوارزمية، وتمثل انعكاسا لعنصرية المجتمع وتحيزه ضد فئة معينة أو دين معين أو طبقة بعينها، أو قد يكون نتيجة تحيز في بيانات التدريب التي يتم تغذية نظام الذكاء الاصطناعي بها"<sup>(3)</sup>.

حيث يمكن تحديد معنى التحيز الخوارزمي وجود انحياز أو تفصيل في خوارزميات المستخدمة في أنظمة الذكاء الاصطناعي سواء في النتائج المتوصل إليها أو تحيز أثناء فترة تعليم الآلة، فتظهر نتيجة التحيز لحساب طائفة على أخرى، وهذا بسبب البيانات المتحيزة التي تم اعتمادها أثناء مرحلة التعلم الآلي. كما يمكن القول كذلك، إن هذه البيانات المتحيزة ما هي إلا امتداد لتحيز وعنصرية البشر التي علمت الآلة وادخلت إليها معلومات ذات طبيعة تحيزية.

أما الكلمة الثانية هي "الخوارزمي"، وتتسبب في الأصل إلى عالم الرياضيات المسلم محمد بن موسى الخوارزمي<sup>(4)</sup>، وتدل "الخوارزمية" لغةً على مجموعة من الخطوات الرياضية والمنطقية والمتسلسلة

---

(1) الخوارزمية: مجموعة من التعليمات المحددة لحل مشكلة، أو أداء مهمة معينة "معجم البيانات والذكاء الاصطناعي إنجليزي- عربي، ط 1، الهيئة السعودية للبيانات والذكاء الاصطناعي، مجمع الملك سلمان العالمي للغة العربية، 2022م، ص 40

(2) المرجع السابق 41.

(3) تحيز الذكاء الاصطناعي... تقنيات وأدوات، مقال منشور على الموقع:

<https://masaar.net/ar> بتاريخ ٢٠٢٣/٣/٧ م، تاريخ الاطلاع ٢٠٢٣/٧/٨ م

(4) كان الباحث الإسلامي في العصور الوسطى أبو جعفر محمد بن موسى الخوارزمي، الذي نسبت إليه الخوارزمية، أكثر اهتمامًا بإجراءات محاسبية تتم خطوة بخطوة موثوقة لحلول الحوسبة للمعادلات الجبرية، وعلى مر العصور تحول مفهوم الخطوات الموثوقة إلى الوظائف الحسابية **Computability and Computable Functions** لإضفاء الطابع المعتمد لفكرة الخوارزمية **Algorithm** وأخي الرّ وصل المفهوم إلى سلسلة محدودة من الإرشادات الدقيقة القابلة للتطبيق على أنظمة الكمبيوتر

Computing systems .Arndt, A. B., "AI-Khwarizmi," Mathematics Teacher, Vol. 76, No. 9, 1983.pp. 668-670 .

Moffat; Algorithmic risk governance: Big data analytics, race and information activism in criminal –Kelly Hannah )5(justice debates, Theoretical Criminology, Vol. 23, I.4, November 2019, p.457.

اللازمة لحل مشكلة ما<sup>(١)</sup>، والمُعَدَّة برمجياً لكي تُعطي نتيجة مُعينة اعتماداً على مُعطيات ومدخلات غُذيت بها<sup>(٢)</sup>. والخوارزميات هي قوام الذكاء الاصطناعي وأهم عناصره، ولا يقتصر دور الخوارزمية فقط على فرز ومزج وتكسير الأرقام من خلال النماذج الرياضية الثابتة ولكنها تقوم بتحديث سلوكها بشكل تكراري استناداً إلى نماذج تم ضبطها استجابة لمجموعة بيانات الإدخال ومقاييس الأداء<sup>(٣)</sup>.

لا يخلو التعريف اللغوي للتحيز الخوارزمي "Algorithmic bias" من أهمية تتعكس بطبيعة الحال على معناه الاصطلاحي، الذي يراه البعض<sup>(٤)</sup> بأنه "التمييز أو التفضيل غير المنصف الذي يظهر في النتائج أو القرارات التي تتخذها الخوارزميات الحاسوبية أو نظم الذكاء الاصطناعي والتي من شأنها التفرقة بين الأشخاص بناءً على عوامل معينة مثل الجنس العرق الدين الأصل الاجتماعي، أو أي سمات أخرى". ويعرفه البعض الآخر<sup>(٥)</sup> بأنه "ظاهرة تتعلق بالتمييز أو الإحجام أو التفضيل غير المبرر يتمثل في تفاوتات غير عادلة أو تمييز غير عادل .

---

(1) intelligence in Our Image The Risks Of Bias And Errors In Artificial Osonde A. Osoba, William Welser Tv, A Intelligence, Rand Corporation, Santa Monice, Calif, 2017, p.4.

(2) د. عبد الله موسي، د. أحمد حبيب بلال، الذكاء الاصطناعي، ثورة في تقنيات العصر، المجموعة العربية للتدريب والنشر، الطبعة الأولى، ٢٠١٩، ص ٩.

(3) Valiant, Leslie, Probably Approximately Correct: Nature's Algorithms for Learning and Prospering in a Complex (8) World, New York: Basic Books, 2013.

(4) Shahriar Akter, Yogesh K. Dwivedi, Shahriar Sajib, Kumar Biswas, Ruwan J. Bandara, Katina Michael Algorithmic bias in machine learning-based marketing models, Journal of Business Research, Vol. 144, 2022,

(5) To stop algorithmic bias, we first have to define it, Emily Bembeneck, Rebecca Nissan, Ziad Obermeyer )10(Brookings, 2021, Accessed 20-8-2023, On this website;

(Brookings, 2021, Accessed 20-8-2023, On this website; /it-define-to-have-first-we-bias-algorithmic-stop-https://www.brookings.edu/articles

## الفرع الثاني

### أسباب التحيز الخوارزمي

#### في أنظمة الذكاء الاصطناعي

يرجع العامل المؤثر والفعال في وجود التحيز الخوارزمي في أنظمة الذكاء الاصطناعي يرجع بالدرجة الأولى إلى التحيز الموجود لدى المبرمجين أثناء تغذية أو تعليم الآلة، ومع هذا فإنه هناك أسباب أخرى أدت إلى تغذية خوارزميات الذكاء الاصطناعي بالتحيز، وبينها كالات (١)

أولاً . البيانات المدخلة ذات الصفات التحيزية:

أن الإنسان هو المعلم الاول والمنشأ للتحيز الخوارزمي الذي أنشأه بناء على وأفكاره الخاصة واتجاهاته وميوله السياسية والثقافية، التي من شأنها التأثير في عمل أنظمة الذكاء الاصطناعي التي تجعل من نتائج تلك الخوارزميات متحيزة لفئة معين دون الأخرى (٢)، وتأكيداً لما سبق يقول خبير الذكاء الاصطناعي تريسان جرين الأمريكي: "إن الآلات بطبيعتها لا يوجد لها انحياز، كما أن الذكاء الاصطناعي لا يريد من الشيء أن يكون صحيحاً أو خطأ لأسباب لا يمكن تفسيرها من خلال المنطق، ولأسف يوجد تحيز بشري في التعلم الآلي بسبب إنشاء خوارزميات البيانات، وحتى الآن لم يحاول أحد حل هذه المشكلة الضخمة... وحيث أن الانحياز صفة بشرية، ولأن الذكاء الاصطناعي يحاكي الذكاء البشري عن طريق التعلم من البيانات المدخلة له فسيرت منه عيوبه كذلك، فإذا احتوت البيانات على عنصرية أو طبقية أو تمييز جنسي، سنكسر الأنظمة هذه الأفكار وتعزز منها(٣)

وترى الكاتبة نوبل أن العنصرية الرقمية ناتجة عن عنصرية المبرمجين الذين يعكسون أفكارهم وآرائهم على تصاميمهم، ما يعني أن مقترحات هذه المحركات ماهي إلا تمثيل بشري لتحيز المطورين وجهلهم بالعواقب الاجتماعية السلبية لتقنياتهم، في النهاية هم من صنعوا أنظمة الخوارزميات(٤)؛ وكمثال عن البيانات المتحيزة التي تظهر في نتائج البحث عن كلمة لاعب، فإن أغلب النتائج ستظهر صورا متعلقة بكرة القدم، في حين قد لا نجد صورة واحدة للاعب السلة او البيسبول، وهذا مرتبط بنوعية البيانات التي تعلم منها، والتي في الأغلب تربط كلمة لاعب بلعبة كرة القدم(٥).

(١) محمد فتحي: الذكاء الاصطناعي ينافس الإنسان في الانحياز مقال منشور على الموقع:

<https://www.alkhaleej.ae> بتاريخ 25 يوليو 2018م، تاريخ الاطلاع 2023/07/08م.

(٢) دان رايلي: للتغلب على التحيز في الذكاء الاصطناعي يجب أولاً الاعتراف بوجوده، منشور على الموقع:

بتاريخ ٢٠٢٣/٧/٤، تاريخ الاطلاع ٢٠٢٤/٧/١٠ <https://fortunearabia.com>

(٣) محمد فتحي: الذكاء الاصطناعي.. ينافس الإنسان في الانحياز، المرجع السابق.

(٤) نور علوان، حفصة جودة: تحيز العالم الرقمي.... عندما تشجعنا محركات البحث على العنصرية،

<https://www.noonpost.com> مقال منشور بتاريخ 2018/10/11م، تاريخ الاطلاع 2023/07/12م.

(٥) محمد فتحي: الذكاء الاصطناعي... ينافس الإنسان في الانحياز، مرجع سابق، تاريخ الاطلاع 2023/07/12م.

كذلك قصة الباحثة في معهد ماساتشوستس للتكنولوجيا جوي بولاميني ذات البشرة السوداء، التي اكتشفت بالصدفة وهي تجرب برنامجاً للتعرف على الوجه أنه لا يتعرف على وجهها، في حين أنه يتعرف بشكل جيد على وجوه أصدقائها أصحاب البشرة البيضاء، فقامت بتجربة قناع أبيض ليتعرف البرنامج على وجهها، وعندها تعرف عليها على الفور<sup>(١)</sup>.

وفي مجال الوظائف وحقوق الإنسان فيما يتعلق بمبدأ المساواة يرجع السبب في تمييز أو تفضيل أفراد معينين من المتقدمين على حساب الآخرين إلى وجود تحيز في البيانات المستخدمة لتدريب الخوارزمية التي تحتوي على تحيزات سابقة، كوجود معايير توظيفية تم الأخذ بها سابقاً تعكس تحيزاً معيناً، وبالتالي، فإن الخوارزمية مع مرور الوقت تتعلم هذا التحيز وتعكسه في عملية انتقاء ملفات الباحثين عن العمل أو المساواة أمام القانون فيما يتعلق بخوارزميات العدالة الجنائية.

### ثانياً: الخوارزميات المتحيزة:

يمكن أن تؤدي الخوارزميات المعيبة إلى تضخيم التحيز من خلال حلقات التغذية الراجعة التي منشأها. فلنتأمل حالة الأنظمة المدربة إحصائياً، مثل تطبيق الترجمة Google Translate، التي تستخدم الضمائر المذكورة، وليس المؤنثة، بشكل افتراضي.

ويرجع هذا التمييز إلى أن نسبة الضمائر المذكورة تبلغ ضعف الضمائر المؤنثة في مصادر نصوص اللغة الإنجليزية. والأسوأ من ذلك أنه في كل مرة يرتكب تطبيق الترجمة هذا الخطأ، يزداد التكرار النسبي للضمير المذكور على شبكة الإنترنت، وهو ما قد يؤدي إلى تقويض التقدم المتحقق بصعوبة نحو تحقيق الإنصاف، والجدير بالذكر أن نسبة الضمائر المذكورة إلى الضمائر المؤنثة قد انخفضت من الأربعة أضعاف في ستينيات القرن الماضي بفضل التحولات الاجتماعية واسعة النطاق<sup>(٢)</sup>.

من أمثلة الخوارزميات المعيبة في مجال العدالة الجنائية التنبؤية ان استخدام خوارزميات معيبة يتعارض مع بعض المبادئ الأصولية للإجراءات الجنائية ولاسيما مبدأ المساواة امام القضاء حيث من حق المتهم في العلم بالأدلة المقامة ضده وحق الوصول إلى كافة المعلومات، والتي كشفت عنه وأظهرته للعيان قضية لوميس ضد ولاية ويسكونسن الأمريكية. والتي أشار فيها أحد القضاة إلى المصالح الخاصة للشركة المطورة لخوارزمية كومباس والمستخدمه في تقييم المخاطر في هذه القضية، بما صرح

---

(١) سكينه السمرة: الذكاء الاصطناعي بين جدلية التفوق والتحيز... هل السبب في الخوارزميات، مقال منشور على الموقع:

<https://www.annahar.com/arabic/section/4> بتاريخ 2023/04/04م، تاريخ الاطلاع 2023/07/15م.

(٢) تصميم الذكاء الاصطناعي بحيث يكون منصفاً مقال منشور على الموقع:

[/https://arabicedition.nature.com](https://arabicedition.nature.com)

بتاريخ 2019/11/26م، تاريخ الاطلاع 2023/07/15م.

به من أنه "الدى Northpointe مصلحة مالية وملكية واضحة في الاستخدام المستمر لخوارزمية COMPAS" وهو ما دفع الشركة إلى الرد بأنه "ال يوجد خلطة سرية لما نقوم به؛ أنها فقط ليست مفهومة بشكل واضح".

وعلى الرغم من هذه الردود المتبادلة، يمتنع على المتهم أو الجمهور - بالفعل - بالوصول الكامل إلى الكود المصدري "الصيغة الجريمة المستقبلية" التي تظهرها هذه الخوارزمية، أو للعلم بكيفية تحديد درجات المخاطر موازنة (١)

ومن أمثلة البرمجة المعيبة للخوارزميات في مجال العمل، قيام المبرمج ببرجمة الخوارزمية مثلاً بـمعايير توظيفية تحيزه كـمعيار أن أصحاب البشرة السوداء لا يـليقون للعمل في الإدارات، أو أن العنصر النسوي لا يصلح للعمل في المصانع، وغيرها من المعايير التحيزية، التي فيها تأثير سلبي على حق الفرد الباحث عن فرصة عمل يضمن بها كرامته في المجتمع.

### ثالثاً: الخوارزميات التي تتصف بـتعميم

يعد تعميم المخاطر وبتها داخل خوارزميات الذكاء الاصطناعي من الأسباب البارزة المؤدية إلى التحيز، فعندما يتم تعميم فكرة أن معظم ذوي البشرة السوداء أناس خطيرون أو مدمني المخدرات أو ذوي سجل إجرامي، أو أن أصحاب اللحي ذوو فكر إرهابي، في هذه الحالة سنقوم الخوارزميات بإطلاق تحذيرات خاطئة ضد أناس أبرياء (٢).

وقد حلّلت وكالة بلومبيرغ أكثر من 5 آلاف صورة أنشأتها شركة ستيبيليتي إيه آي، ووجدت أن برنامجها ضخم الصور النمطية حول العرق والجنس، إذ يصور عادةً الأشخاص ذوي البشرة الفاتحة على أنهم يشغلون وظائف ذات رواتب عالية، بينما صنّف الأشخاص ذوي البشرة الداكنة على أنهم "يعملون في غسيل الأطباق" و"حراس أمن منازل" (٣).

## الفرع الثالث

### أنواع التحيز الخوارزمي

يتجسد التحيز الخوارزمي في أنظمة الذكاء الاصطناعي على ثلاثة أنواع هي كالآتي (٤):

(١) خوارزميات الذكاء الاصطناعي والعدالة الجنائية التنبؤية دراسة وصفية تحليلية تأصيلية مقارنة دكتور/ طارق أحمد ماهر زغول استاذ القانون الجنائي المساعد كلية الحقوق - جامعة عين شمس ص ٤٥ .

(٢) طه الراوي: عنصرية التكنولوجيا.... عن أسباب وأشكال التحيز الرقمي، مقال منشور على الموقع:

بتاريخ 2020/06/12، تاريخ الاطلاع 2023/07/15م

<https://www.noonpost.com/content/37283>

(٣) الذكاء الاصطناعي متهم بالعنصرية، مقال منشور على الموقع

<https://www.alaraby.co.uk> بتاريخ 2023/06/06م، تاريخ الاطلاع 2023/07/16م.

(٤) مصطفى السيد حسين: عنصرية الذكاء الاصطناعي.... لماذا يبدو أصدقاؤك بـملاح أوروبية عند

## ١- التحيز بسبب التفاعل:

يعد التحيز التفاعلي أحد أشكال التحيز التي تحدث للذكاء الاصطناعي بسبب تفاعل المستخدمين، فعلى سبيل المثال هناك تطبيقات تقوم بتحويل الصورة إلى صورة جنس آخر، ولو كان غالبية مستخدمي هذه التطبيقات من أصحاب البشرة البيضاء، فإنه مع مرور الوقت ستقدم تصورات أقرب للمستخدمين البيض، بسبب تفاعلهم الكبير مع هذه التطبيقات، كصورة يقوم المستخدم صاحب البشرة البيضاء برفعها إلى التطبيق، وهو ما يؤدي إلى خلق نوع من العنصرية أثناء استخدام أنظمة الذكاء الاصطناعي.

## ٢- التحيز الكامن:

يعتبر الانحياز الكامن أحد أهم أسباب انحياز الذكاء الاصطناعي، حيث يقوم هذا الأخير في العادة بدراسة مجموعة من البيانات لإيجاد روابط بينها والتعلم منها، في حين تكون هذه البيانات في الغالب منحازة إلى ثقافة شعب من الشعوب، ومثال على ذلك احتلال اللغة الانجليزية 55 بالمئة من الانترنت، بينما اللغة العربية تحتل 6.1 بالمئة، وتأسيسا على ذلك، فإن محرك جوجل للترجمة القائم على الذكاء الاصطناعي ستكون ترجمته باللغة الانكليزية أفضل وأدق من اللغة العربية

## ٣- التحيز الاختياري:

وهو التحيز في شكل البيانات التي يختارها المبرمج أو صانع التطبيق لتدريب الذكاء الاصطناعي، فمثلا لو أراد المبرمج تدريب تطبيقه أولاً على تمييز الجنس إن كان رجل أو امرأة، وثانياً على الصور الشخصية على ملفات تويتر، فإن الولايات المتحدة الأمريكية ستكون في الصدارة من ناحية عدد مستخدمي تويتر، وبالتالي فإن اختيار المبرمج لهذه البيانات لتحديد لتدريب الذكاء الاصطناعي خلق انحيازاً في الخوارزمية.

لا شك أن التحيز الخوارزمي نوع من أنواع التمييز العنصري، الذي قد يتحقق بإحدى صورتيه التمييز المباشر وغير المباشر. ويعد التمييز مباشراً عندما يعامل الشخص في وضع ما بشكل أقل تفضيلاً بسبب جنسه وأصله وحالته الصحية أو لغير ذلك من الأسباب، مقارنة بشخص آخر لا يكون كذلك، أو كان أو سيكون في حالة مماثلة أما التمييز غير المباشر هو تدبير محايد، لكنه يلحق الضرر بشكل خاص بمجموعة من الأشخاص على أساس معيار غير قانوني للتمييز، والذي لا يُبرره موضوعياً هدف مشروع، مع اتسام وسائل تحقيق الهدف بعدم الضرورة والتناسب.

## الفرع الرابع

### اركان وصور الأخلال بالحق محل الحماية

المشروع المصري أعطى قيمة دستورية لمبادئ المساواة والمواطنة وعدم التمييز<sup>(١)</sup> حيث نصت المادة ٤ من الدستور على أن "السيادة للشعب وحده، يمارسها ويحميها، وهو مصدر السلطات، ويصون

<sup>١</sup> - كما نصت المادة ٩ على أنه تلتزم الدولة بتحقيق تكافؤ الفرص بين جميع المواطنين، دون تمييز".

وحدته الوطنية التي تقوم على مبادئ المساواة والعدل وتكافؤ الفرص بين جميع المواطنين، وذلك على الوجه المبين في الدستور".، بموجب ما نص عليه في الدستور المصري لعام ٢٠١٤ والمعدل في ٢٠١٩ من تجريم التمييز والحض على الكراهية، والتمييز الذي أشار له الدستور لا يقتصر على الصور التي عددها الدستور وإنما يشمل كافة أنواع التمييز الأخرى لا سيما التحيز الخوارزمي، ولعل ذلك يُستشف من أحد أهم أحكام المحكمة الدستورية العليا بقولها "لئن نص الدستور على حظر التمييز بين المواطنين في أحوال بعينها هي تلك التي يقوم التمييز فيها على أساس من الأصل أو الجنس أو اللغة أو الدين أو العقيدة، إلا أن إيراد الدستور لصور بذاتها يكون التمييز فيها محظوراً، مرده أنها الأكثر شيوعاً في الحياة العملية، ولا يدل البتة على انحصاره فيها، إذ لو صح ذلك - وهو غير صحيح - لكان التمييز فيما عداها جائزاً دستورياً، وهو ما يناهض المساواة التي كفلها الدستور، وينقض أسسها ويعطل مقاصدها وآية ذلك أن من صور التمييز التي أغفلتها المادة ٤٠ من الدستور ما لا تقل عن غيرها وزناً وخطراً سواء في محتواها، أو من جهة الآثار التي تتولد عنها وترتبتها كالتمييز بين المواطنين في نطاق حقوقهم وحررياتهم لاعتبار مرده إلى المولد أو الثروة أو المركز الاجتماعي أو انتمائهم الطبقي أو ميولهم الحزبية وآرائهم، أو عصابيتهم القبلية، أو نزعاتهم العرقية، أو إلى موقفهم من السلطة العامة وإعراضهم عن تنظيماتها، أو مناوأتهم لها، أو تبنيهم لأعمال بذاتها<sup>(١)</sup>... كذلك فإن صور التمييز التي تناقض مبدأ المساواة أمام القانون، وإن تعذر حصرها<sup>(٢)</sup>، إلا أن قوامها كل تفرقة أو تقييد أو تفضيل أو استبعاد، ينال بصورة تحكمية من الحقوق والحرريات التي كفلها الدستور أو القانون، وذلك سواء بإنكار أصل وجودها، أو من خلال تقييد آثارها بما يحول دون مباشرتها على قدم من المساواة الكاملة بين المؤهلين قانوناً للانتفاع بها"<sup>(٣)</sup>

كذلك نصت الفقرة الأولى من المادة ١١ على أن تكفل الدولة تحقيق المساواة بين المرأة والرجل في جميع الحقوق المدنية والسياسية والاقتصادية والاجتماعية والثقافية وفقاً لأحكام الدستور".

١ - نصت المادة ٥٣ من الدستور على أن المواطنون لدى القانون سواء، وهم متساوون في الحقوق والحرريات والواجبات العامة، لا تمييز بينهم بسبب الدين، أو العقيدة، أو الجنس، أو الأصل، أو العرق، أو اللون، أو اللغة أو الإعاقة، أو المستوى الاجتماعي، أو الانتماء السياسي أو الجغرافي لأي سبب آخر التمييز والحض على الكراهية جريمة يعاقب عليها العالم تلتزم الدولة باتخاذ التدابير اللازمة للقضاء على كافة أشكال التمييز، وينظم القانون إنشاء مفوضية مستقلة لهذا الغرض.

٢ - حكم المحكمة الدستورية العليا، القضية رقم ٤٠ لسنة ١٦ قضائية دستورية، بتاريخ جلسة ٢ من سبتمبر ١٩٩٥.

٣ - في سابقة قضائية حديثة للمحكمة الدستورية العليا بالجلسة العلنية المنعقدة بـ يوم السبت الثاني من سبتمبر سنة ٢٠٢٣م، الموافق السابع عشر من صفر سنة ١٤٤٥ هـ، في الدعوى المقيدة بجدول المحكمة الدستورية العليا برقم ٧٠ لسنة ٤٣ قضائية "دستورية"، قضت بانه: "وحيث أن الدستور قد اعتمد بمقتضى نص المادة (٤). منه مبدأ المساواة، باعتباره - إلى جانب مبدأي العدل وتكافؤ الفرص - أساساً لبناء المجتمع وصون وحدته الوطنية، وتأكيداً

هذا التجريم الدستوري للتمييز بين الأشخاص انعكس أثره بالتبعية على قانون العقوبات المصري الذي جرم التمييز وحدد المشرع أركان جريمة التحيز تفصيلاً بالمادة ١٦١ مكرراً والتي تنص على أنه " يعاقب بالحبس وبغرامة لا تقل عن ثلاثين ألف جنيه ولا تجاوز خمسين ألف جنيه أو بإحدى هاتين العقوبتين كل من قام بعمل أو بالامتناع عن عمل يكون من شأنه إحداث التمييز بين الأفراد أو ضد طائفة من طوائف الناس بسبب الجنس أو الأصل أو اللغة أو الدين أو العقيدة وترتب على هذا التمييز إهدار المبدأ تكافؤ الفرص أو العدالة الاجتماعية أو تكدير للسلم العام .

حيث حصر المشرع أركان جريمة التحيز في الفعل أو الامتناع بقوله..... قام بعمل أو بالامتناع عن عمل يكون من شأنه إحداث التمييز بين الأفراد أو ضد طائفة من طوائف الناس بسبب الجنس أو الأصل أو اللغة أو الدين أو العقيدة وترتب على هذا التمييز إهدار المبدأ تكافؤ الفرص أو العدالة الاجتماعية أو تكدير للسلم العام؛ أما عن الركن المعنوي فالعلم و الإرادة المقترضين في أي من الجرائم القصد العام .

وبالنظر إلى السلوك المكون للركن المادي الجريمة التمييز فإنه يشمل كل عمل أو الامتناع عن عمل يكون من شأنه إحداث التمييز بين الأفراد أو ضد طائفة من طوائف الناس بسبب الجنس أو الأصل أو اللغة أو الدين أو العقيدة وترتب على هذا التمييز إهدار المبدأ تكافؤ الفرص أو العدالة الاجتماعية أو تكدير للسلم العام ورغم أن المشرع لم يصرح بأن التحيز الخوارزمي احد أنماط التمييز، إلا أنه بلا شك يدخل ضمن الأفعال المكونة لسلوك الركن المادي لأن من شأن هذا التحيز إحداث التمييز بين الأفراد بسبب سمة بشرية معينة كالجنس أو الأصل أو اللغة أو الدين أو العقيدة أو غيرها من أسباب التمييز الأخرى لأن المشرع عدد هذه الأسباب على سبيل المثال لا الحصر.

وشددت الفقرة الثانية من هذه المادة العقوبة إذا وقعت الجريمة من موظف عام أو مستخدم عمومي أو أي إنسان مكلف بأداء خدمة عامة، فنصت على أن وتكون العقوبة الحبس مدة لا تقل عن ثلاثة أشهر والغرامة التي لا تقل عن خمسين ألف جنيه ولا تجاوز مائة ألف جنيه أو بإحدى هاتين العقوبتين

---

لذلك حرص الدستور في المادة (٥٣) منه على كفالة تحقيق المساواة لجميع المواطنين أمام القانون في الحقوق والحريات والواجبات العامة، دون تمييز بينهم لأي سبب، إلا أن ذلك لا يعني - وفقاً لما استقر عليه قضاء هذه المحكمة - أن تعامل فئاتهم على ما بينها من تفاوت في مراكزها القانونية معاملة قانونية متكافئة، كذلك لا يقوم هذا المبدأ على معارضة صور التمييز جميعها، ذلك أن من بينها ما يستند إلى أسس موضوعية ولا ينطوي بالتالي على مخالفة لنص المادتين (٤) و (٥٣) المشار إليهما، بما مؤداه أن التمييز المنهي عنه بموجبهما هو ذلك الذي يكون تحكيمياً، وأساس ذلك أن كل تنظيم تشريعي لا يعتبر مقصوداً لذاته، بل لتحقيق أغراض بعينها تعكس مشروعيتها إطاراً للمصلحة العامة التي يسعى المشرع إلى تحقيقها من وراء هذا التنظيم ."

إذا ارتكبت الجريمة المشار إليها في الفقرة الأولى من هذه المادة من موظف عام أو مستخدم عمومي أو أي إنسان مكلف بخدمة عمومية.

أما عن صور الأخلال بالحق محل الحماية فالتحيز الخوارزمي معاقب عليه لدى المشرع المصري بوصفه أحد صور التمييز العنصري رغم أن هذا السلوك لم ترد بشأنه أي تطبيقات قضائية لندرة حدوثه في مصر لعدم وجود منصات أو محركات بحث أو شركات تكنولوجيا ضخمة على الإنترنت فيها، إلا أن النص سالف الذكر يستغرق هذه الصورة المستحدثة من صور التمييز.

ويرى الباحث أنه من الأفضل إدخال تعديل بالنص صراحة على تجريم التحيز الخوارزمي في فقرة ثالثة لنص المادة ١٦١ مكرراً مع تشديد العقوبة على هذا السلوك، وذلك لأن الخطورة الإجرامية للتحيز الخوارزمي أكثر جسامة من صور التمييز العنصري الأخرى وذلك لما تتسم به هذه الصورة المستحدثة من سرية ممنهجة في التمييز بين الأفراد وهو ما يتعارض مع مبدأ الشفافية والعدالة الخوارزمية، فصعوبة الكشف عن التحيز الخوارزمي هي إحدى مخاطره، لأن الخوارزميات غالباً ما تكون سرية، ولا يتم الإفصاح عن تفاصيلها إلا للمبرمجين والخبراء، كما أن الخوارزميات تكون معقدة للغاية، ما يجعل من الصعب فهم كيفية عملها وكيفية الكشف عن التحيزات المحتملة فيها وهو ما يعزز تضمين المادة بظرف مشدد ينص عليه صراحة حال استخدام خوارزميات متحيزة عن علم، أو في الحالة التي ترفض فيها الشركة تصحيح الخوارزميات المتحيزة خلال فترة زمنية محددة إذا لم تكن على علم بها.

وبخصوص الموامة التشريعية بين القانون المصري وبعض المواثيق الدولية، فقد علق المجلس القومي لحقوق الإنسان على مدي تنفيذ الحكومة المصرية للتوصيات الختامية للجنة القضاء على التمييز العنصري بتاريخ ٢٠٠١/٩/١٥ بأنه على الرغم من النصوص الدستورية الجيدة سابق الإشارة إليها صور التحيز الخوارزمي نوع من أنواع التمييز العنصري، الذي قد يتحقق بإحدى صورتيه التمييز المباشر وغير المباشر وبعد التمييز مباشراً عندما يعامل الشخص في وضع ما يشكل واصله وحالته الصحية أو لغير ذلك من الأسباب، مقارنة بشخص آخر لا يكون كذلك، أو كان أو سيكون في حالة ممثلة<sup>(١)</sup>. أما التمييز غير المباشر هو تدبير محايد، لكنه يلحق الضرر بشكل خاص بمجموعة من الأشخاص على أساس معيار غير قانوني للتمييز، والذي لا يبرره موضوعياً هدف مشروع، مع أقسام وسائل تحقيق الهدف بعدم الضرورة والتناسب<sup>(٢)</sup>.

(١) تصميم الخوارزمية ذاتها يمكن أن يكون تصميم الخوارزمية نفسه مصدر للتمييز، سواء عن طريق استخدام متغيرات غير معالمة أو نماذج رياضية تحتوي على افتراضات محددة تؤدي إلى التمييز السابق الاجتماعي والثقافي يمكن أن يعكس التغير في الخوارزميات عن طريق العكاس القيم والمعتقدات المجتمعية والثقافية التي يتم تدريب الخوارزمية عليها.

(٢) السابق الاجتماعي والثقافي يمكن أن يعكس التغير في الخوارزميات عن طريق انعكاس القيم والمعتقدات المجتمعية والثقافية التي يتم تدريب الخوارزمية عليها.

وعلى خلاف ما هو معتقد بأن الخوارزميات تتمتع بالحياد المطلق<sup>(١)</sup>، إلا أنها في حقيقة الأمر تنتج هذين النوعين من التمييز استناداً إلى آلية عمل الخوارزمية من خلال تحليل البيانات واستخلاص الروابط والعلاقات بينها والتي قد تكون مشوهة بما تحتويه من تمييز يتم تدريب الخوارزميات عليها، فضلاً عما يشوب الخوارزمية ذاتها من تصميم مشوه غير محايد<sup>(٢)</sup> يؤدي إلى إعادة إنتاج الاتجاهات المتحيزة، بل وتعزيزها<sup>(٣)</sup> بناء عليه يكون التحيز الخوارزمي نمط حديث من أنماط التمييز العنصري ويخضع بالتبعية للنصوص القانونية التي تجرم التمييز بين الأشخاص لا سيما قانون العقوبات المصري في (١٦١) مكرر الذي جرم التمييز بكافة أشكاله، بل والتحرير عليه أيضاً في المادة (١٧٦) <sup>(٤)</sup> منه. والأمر ذاته في قانون العقوبات الأمريكي الفيدرالي في القسم (٢٤٥) منه <sup>(٥)</sup>

Galiskan, A., Bryson, J. J. & Narayanan, A ..

تحتوي الدلالات المشتقة تلقائياً من مجموعات اللغة على تحيزات شبيهة بالتحيزات البشرية، مجلة ساينس، المجلد ٣٥٦ . ١٦٣٣٤ . ٢٠١٧ . ص ١٨٣-١٨٦ . تاريخ الوصول ٢٠-٨-٢٠٢٣ . على هذا الموقع:

<https://www.science.org/doi/10.1126/science.aal4230>

<sup>(١)</sup>[https://www.theguardian.com/artanddesign/2023/aug/12/rafael-lozano-hemmers-atmospheric-memory-exhibition-powerhouse-museum-sydney-things-to-do-weekend?utm\\_source=ground.news&utm\\_medium=referral](https://www.theguardian.com/artanddesign/2023/aug/12/rafael-lozano-hemmers-atmospheric-memory-exhibition-powerhouse-museum-sydney-things-to-do-weekend?utm_source=ground.news&utm_medium=referral)

<sup>(٢)</sup> طارق أحمد ماهر زغلول، خوارزميات الذكاء الاصطناعي والعدالة الجنائية التنبؤية، المرجع السابق، ص ٢٢١ تنص المادة (١٦١) مكرراً) من قانون العقوبات المصري المضافة بالقانون رقم ١٢٦ لسنة ٢٠١١، على أن يعاقب بالحبس وبغرامة لا تقل عن ثلاثين ألف جنيه ولا تجاوز خمسين ألف جنيه أو بإحدى هاتين العقوبتين كل من قام بعمل أو بالامتناع عن عمل يكون من شأنه إحداث التمييز بين الأفراد أو ضد طائفة من طوائف الناس بسبب الجنس أو الأصل أو اللغة أو الدين أو المفيدة وترتب على هذا التمييز إصدار لمبدأ تكافل الفرص أو العدالة الاجتماعية أو تكبير السلم العام، وتكون العقوبة الحبس مدة لا تقل عن ثلاثة أشهر والغرامة التي لا تقل عن خمسين ألف جنيه ولا تجاوز مائة ألف جنيه أو بإحدى هاتين العقوبتين إذا ارتكبت الجريمة المشار إليها في الفقرة الأولى من هذه المادة من موظف عام أو مستخدم عمومي أو أي إنسان مكلف بخدمة عمومية". نص المادة (١٧٦) من قانون العقوبات =المصري المحلة بالقانون رقم ١٩٧ لسنة ٢٠٠٦ على أن يعاقب بالحبس كل من حرض بإحدى الطرق المتقدم ذكرها على التمييز ضد طائفة من طوائف الناس بسبب الجنس أو الأصل أو اللغة أو الدين أو العقيدة إذا كان من شأن هذا التحريض تكبير السلم العام

<sup>(٣)</sup> المرجع السابق ص ٦٥ .

<sup>(٤)</sup> تنص المادة (١٧٦) من قانون العقوبات المصري المعدلة بالقانون رقم ١٤٧ لسنة ٢٠٠٦ على أن: " يعاقب بالحبس كل من حرض بإحدى الطرق المتقدم ذكرها على التمييز ضد طائفة من طوائف الناس بسبب الجنس أو الأصل أو اللغة أو الدين أو العقيدة إذا كان من شأن هذا التحريض تكبير السلم العام."

<sup>(٥)</sup> Her or not acting under color of law, Federally protected activities; (b)Whoever, whet -U.S. Code § 245 18 )5(by force or threat of force willfully injures, intimidates or interferes with, or attempts to injure, intimidate or interfere with— ....

## المطلب الثاني

### التحيز أثر تغذية خوارزميات الذكاء الاصطناعي

جاءت تقنية الذكاء الاصطناعي بصور جديدة من انتهاكات حقوق الإنسان، وبأشكال مستحدثة من الاضطهاد الممارس على الأفراد وعادة ما يكون ضحايا هذه الممارسات من الأشخاص الهشين والمهمشين الذين هم أصلا في أمس الحاجة إلى الحماية القانونية لحقوقهم الأساسية خاصة الحق في العمل، من الانتهاك والاضطهاد؛ فالاضطهاد والتهميش لهذه المجموعات ينعكس في البيانات. في سبيل بيان ذلك سوف يقسم الباحث هذا المبحث الى فرعين أثر تغذية خوارزميات الذكاء الاصطناعي على قطاع العدالة وحق التوظيف باعتبارها حق من الحقوق الأساسية للإنسان .

### الفرع الأول

#### أثر التحيز الخوارزمي على المحاكمات الجنائية

حيث أضحى استخدام الخوارزميات في أنظمة العدالة الجنائية يثار مخاوف جدية فيما يتعلق بالمادة السادسة المتعلقة بالحق في محاكمة عادلة من الاتفاقية الأوروبية لحقوق الإنسان<sup>(1)</sup>، والمادة ٤٧ من ميثاق الحقوق الأساسية لاتحاد الأوروبي، ومبدأ المساواة في القانون وإجراءات الدعوى الجنائية كما أقرتها المحكمة الأوروبية لحقوق الإنسان<sup>(2)</sup>، وتضمن معايير المحاكمة العادلة الواردة في المادة السادسة من الاتفاقية الأوروبية لحقوق الإنسان<sup>(3)</sup> للمتهمين الحق في المشاركة الفعالة في المحاكمة، وتشمل افتراض البراءة، والحق في معرفة سبب الاتهام وطبيعته، والحق في الدفاع عن النفس... إلخ.

وتجدر بنا الإشارة إلى أن نظام حماية البيانات الشخصية غير كافٍ لمعالجة جميع التحديات فيما يتعلق بضمان امتثال أنظمة الذكاء الاصطناعي لحقوق الإنسان، ومن ثم، فإن الآثار المترتبة على حقوق الإنسان متعددة، ومن أهم حقوق الإنسان التي قد تتأثر من خلال استخدام خوارزميات المعالجة الآلية هي<sup>(3)</sup>: الحق في محاكمة عادلة ومراعاة الأصول القانونية، والخصوصية وحماية البيانات، وكذلك حرية التعبير والحق في المساواة وحظر التمييز، إضافة إلى الحقوق الاجتماعية والحصول على الخدمات العامة، ونظرًا؛ لأن الحريات الأساسية مترابطة ومتشابكة من المحتمل أن تتأثر جميع حقوق الإنسان باستخدام الخوارزميات، على سبيل المثال الرعاية الاجتماعية والأنظمة القضائية، لذلك قد يؤدي تطور

---

(1) Convention for the Protection of Human Rights and Fundamental Freedoms (European Convention on Human Rights(here-inafter: ECHR)

(2)Charter of Fundamental Rights of the European Union [2012] OJ C 326 (hereinafter: Charter).

(3) Alexandra Chouldechova, Fair prediction with disparate impact: A study of bias in recidivism prediction instruments,

[https://www.andrew.cmu.edu/user/achoulde/files/disperate\\_impact.pdf](https://www.andrew.cmu.edu/user/achoulde/files/disperate_impact.pdf)

"خوارزميات الذكاء الاصطناعي" المستخدمة في النظم والمجالات القانونية إلى تعطيل مفهوم حقوق الإنسان كحق دستوري.

وهكذا، يتم انتهاك الحق في المشاركة الفعالة في مجموعة متنوعة من الإجراءات الجنائية، بدءاً من المحاكمة عن بعد عبر شاشات الفيديو في قاعة المحكمة لمنع المتهم من الحضور في المحاكمة أو من مناقشة شاهد يشهد ضد المتهم، وهذا الآخر هو أيضاً أحد الضمانات للمحاكمة العادلة الواردة في الفقرة الثالثة من المادة السادسة - آفة الذكر - ويتطلب تقديم جميع الأدلة ضد المتهم في حضوره وحضور أطراف الدعوى الجنائية في جلسة استماع علنية، مما يمنح المتهم فرصة فعالة لتنفيذ أدلة الإثبات، ولا ينطبق الحق في المواجهة على الشهود فقط، ويشمل أيضاً الخبراء والمجني عليهم، في أي حالة تكون فيها البيانات إلى حد ما كأساس لإدانة الجاني، فإنه يشكل دليلاً على المقاضاة التي تنطبق عليها ضمانات المحاكمة العادلة، يمكن تطبيق الحق المنصوص عليه في المادة ٣/٦ د على الأدلة المستندية وملفات الحاسب الآلي ذات صلة بالاتهامات الجنائية ضد المتهم. لذلك، من أجل ضمان المشاركة الفعالة في المحاكمة، يجب أن يكون المجني عليه قادراً على الرد على الأدلة المؤتمتة التي هي أساس إدانته، ومع ذلك، فإن الحق في المواجهة ليس مطلقاً ويمكن تقييده إذا تم استيفاء شروط معينة.

وعليه، فقد كان النهج التقليدي للمحكمة الأوروبية لحقوق الإنسان هو انتهاك الحق في محاكمة عادلة إذا كانت الإدانة تستند على دليل مؤتمت<sup>(١)</sup>، ومع ذلك، ابتعدت المحكمة جزئياً عن اجتهاداتها السابقة، مشيرة إلى أن قبول الأدلة المؤتمتة لن يؤدي تلقائياً إلى خرق المادة ١/٦ عند تقييم نزاهة المحاكمة، وللمحكمة الأوروبية لحقوق الإنسان أن تتظر فيما إذا كان من الضروري قبول مثل هذه الأدلة وما إذا كانت هناك أدلة كافية، بما في ذلك ضمانات إجرائية قوية، إن الإشكالات التي طرحها أنظمة الذكاء الاصطناعي تشبه إلى حد كبير تلك الإفادات التي قدمها شهود مجهولون أو أدلة مستنديه غير مفصح عنها مما يؤدي إلى عدم شفافية أنظمة الذكاء الاصطناعي.

وبناءً عليه، بات من الضروري ضمان إتاحة الفرصة للمتهم لتنفيذ الأدلة المقدمة ضده أو لموازنة عبء عدم الكشف عن هوية الشهود السريين، وفي ظل قواعد صارمة تضمن عدم وضع المتهم في وضع غير مؤاتي، يجب تطبيق مثل هذه القاعدة على استخدام أنظمة الذكاء الاصطناعي المستخدمة في بيئات العدالة الجنائية، لتحقيق توازن عادل بين الحق في المشاركة الفعالة في المحاكمة من جهة، واستخدام أنظمة الذكاء الاصطناعي الغامضة المصممة لمساعدة القضاة على الوصول إلى تقييمات

---

(1) Cour Européenne Des Droits De L'homme, European Court Of Human Rightscase of Doorson V. The Netherlands (Application <https://www.no.0524/92>), Council Of Europe, Judgmentstrasbourg,hrdp.org/files/2013/09/08/CASE\_OF\_DOORSON\_v.\_THE\_NETHERLANDS\_. 26 March 1996, p. 11. pdf

أكثر دقة لسلوك المتهم في المستقبل من جهة أخرى، وينبغي تفسير الحق في استجواب الشهود بحيث يشمل أيضاً الحق في فحص البيانات والقواعد الأساسية لمنهجية تقييم المخاطر، في إجراءات التقييم، ويستتبع مثل هذا الحق التأكد من أنه يمكن للشخص المدان أن يشكك في البيانات التي يتم إدخالها في الخوارزمية. وبالتالي، هل تنتهك "خوارزميات الذكاء الاصطناعي" حق المتهم في الإجراءات القانونية الواجبة؟

حيث إن الصيغة المحددة لتحديد خوارزمية تقييم المخاطر هي ملكية خاصة، فإن الأطراف غير قادرين على الطعن في صحة النتائج، وهذا ينتهك حق المتهمين في المحاكمة العادلة، وتطبيقاً لذلك تم الطعن في استخدام خوارزميات كومباس في إصدار الأحكام في قضية لوميس ضد ويسكونسن<sup>(1)</sup> باعتباره انتهاكاً لحق المتهم في الإجراءات القانونية الواجبة لسببين:

الأول: هو طبيعة الملكية ل COMPAS تمنع المتهم من مناقشة الصلاحية العلمية لتقييم COMPAS.

ثانياً: يجادل لوميس بأن الخوارزمية غير دستورية بسبب الطريقة التي تنظر بها إلى جنس أطراف الدعوى، فلدى COMPAS مقياس منفصل للنساء والرجال<sup>(1)</sup>.

وكذلك ، توصلت محكمة الاستئناف في ولاية كنساس في قضية كنساس ضد والنز "2017"<sup>(2)</sup>، إلى حكم مخالف لخوارزمية "Loomis" حيث قررت المحكمة أنه يجب السماح للمتهم بالوصول إلى الحق الكامل لتقييم مراجعة مخزون بيانات خوارزمية "LSIR"، والذي اعتمدت عليه المحكمة في تقرير شروط الاختبار التي تفرض عليه، وقد قررت محكمة الاستئناف أنه برفض وصول المتهم إلى تقييم LSIR، حرمته محكمة المقاطعة من فرصة الطعن في دقة البيانات التي كان مطلوباً من لمحكمة الاعتماد عليها في تحديد شروط فترة الاختبار القضائي. وكذلك في قضية كنساس ضد إيستلينج<sup>(3)</sup>، قررت محكمة الاستئناف أن فشل المحكمة المحلية في منح المتهم من حق الاطلاع على نسخة من خوارزمية LSIR بالكامل حرمة من حقه الدستوري في الإجراءات القانونية الواجبة في مرحلة إصدار الحكم في إجراءاته الجنائية.

ولذلك، نرى أنه لتطبيق مبدأ حق التقاضي في أجل معقول وضمان تسريع إصدار القرارات والأحكام الجنائية، ولإنجاح هذا التوجه بقطاع العدالة، صار من الواجب صياغة وسن التشريعات التي

---

(1) Megan Garber, When Algorithms Take the Stand, The Atlantic (Jun. 30, 2020), <http://www.theatlantic.com/technology/archive/2016/06/when-algorithms-take-the-stand/48956>

(2) <https://www.brookings.edu/blog/techtank/algorithms-and-sentencing-what-does-due-process-require> (last visited

(3) <http://sbpublishing.org/index.html> (last visited .) 2020/7/6

تحدد النطاق العام الذي تعمل فيه "خوارزميات الذكاء الاصطناعي"، وما قد يترتب عليها من أخطاء، وإدراكاً لأهمية وجود قوانين حديثة ملائمة لتنظم طبيعة عمل التطبيقات القائمة على برمجية الذكاء الاصطناعي، إلى جانب تعميم القواعد والمبادئ التوجيهية لمصممي برامج الخوارزميات للتعامل معها عند الحاجة.

## الفرع الثاني

### أثر التحيز الخوارزمي

#### على الحق في التوظيف

يعد الحق في العمل من الحقوق الأساسية للإنسان، كيف لا، وهو الحق الذي يمكنه من العيش والتمتع بحياة كريمة وهنية، وقد حرصت الشريعة الإسلامية الغراء والقوانين الوضعية على إعطاء كل فرد فرصة العمل في المجال الذي له فيه كفاءة ومؤهلات، محاربة كل تمييز بين الأفراد في الحصول على الوظيفة بناء على العرق أو الجنس أو الدين.

وعلى الرغم من هذا الحرص والتأكيد على حق كل فرد في العمل، فإن ظهور الذكاء الاصطناعي وانتشاره السريع واستخدامه الواسع أدى إلى حدوث انتهاك صارخ لهذا الحق، من قبل شركات التوظيف، ذلك أن أنظمة الذكاء الاصطناعي في حد ذاتها تنطوي على تغيير الطريقة التي تعمل بها الشركات بل الحكومات في العالم، على نحو يجلب احتمالية التدخل الهائل في حقوق الإنسان<sup>(١)</sup>.

حيث يتخذ نظام الذكاء الاصطناعي الذي يتم تدريبه على البيانات من عمليات التوظيف المتحيزة قراراتٍ تمييزية بشأن من ينبغي أن يتم تعيينه، حيث تعقد المؤسسات العزم على استخدام الأدوات التي تعمل بالذكاء الاصطناعي لفحص المرشحين للوظائف واختيارهم بسبب الفوائد التي يقدمها، مثل زيادة الموضوعية والكفاءة، وتثير هذه الأدوات مخاوف بشأن أضرارها المتعلقة بالتحيز والإنصاف<sup>(٢)</sup>.

وكأحد الأمثلة الشهيرة على انتهاك الذكاء الاصطناعي للحق في العمل ما حدث في شركة أمازون التي قامت في عام 2014 ببناء نظام ذكاء اصطناعي للمساعدة في عملية التوظيف بالشركة، لتجد أن النظام الجديد يميز ضد النساء، حيث.

شكّلت الشركة فريقاً من المبرمجين للعمل على جعل عملية اختيار الموظفين أوتوماتيكية، هذا الفريق غذى نظام الذكاء الاصطناعي بمجموعة كبيرة من السير الذاتية التي قُدمت إلى أمازون على مدار 10 سنوات، وكان من المفترض أن يقوم النظام- بناءً على البيانات التي تم تغذيته بها -بتقديم توصيات واختيار أفضل المرشحين للوظائف الجديدة، بعد مرور عام، لاحظ المهندسون أن النظام يميز

(١) هايدي عيسى حسن علي حسن: حقوق الإنسان في عصر الذكاء الاصطناعي "معطيات ورؤى وحلول"، مجلة الشريعة والقانون، ع 85، يناير 2021م، ص 260.

(٢) الذكاء الاصطناعي: تحقيق التوازن بين الابتكار وحقوق الإنسان مقال منشور على بتاريخ ٢٢/٥/٢٠٢٣

ضد النساء، ومع مراجعة السير الذاتية التي تم تغذية النظام بها لتدريبه على اختيار الأنسب للوظائف، اكتشفت الشركة أن أغلب هذه السير الذاتية كانت لذكور، وبالتالي أعطى نظام الذكاء الاصطناعي تقييماً منخفضاً للسير الذاتية التي تحتوي على أسماء مؤنث، أو بيانات أخرى مثل ممارسة رياضة معينة<sup>(1)</sup>.

ومن الأمثلة كذلك، التي يظهر فيها انحياز أنظمة الذكاء الاصطناعي على أساس العرق في مجال التقديم للوظائف، ما حصل في شركة "فوكس نيوز"<sup>(2)</sup>، التي تُستخدم الخوارزمية لاختيار طلبات التوظيف المرموقة بناءً على من كان ناجحاً من الموظفين في فوكس نيوز في الماضي، حيث تركز الخوارزمية على شرطين تحديداً: أن الموظف قد عمل خمس سنوات، وأنه قد ترقى مرتين على الأقل خلال تلك الفترة، وحينها تقوم الخوارزمية بتصفية المتقدمين المقارنين للموظفين القدامى المستوفين لهذين الشرطين غير أن معظم موظفي فوكس نيوز المستوفين لهذين الشرطين في الماضي هم من الرجال البيض (العرق القوقازي)، وكننتيجة غير مفاجئة لذلك، فقد كانت معظم طلبات التوظيف التي اختيرت من قِبَل الخوارزمية على أنها واعدة هي لمتقدمين من نفس العرق<sup>(3)</sup>.

إن لاعتماد الشركات وأرباب العمل على أنظمة التوظيف الآلي القائمة على أنظم الذكاء الاصطناعي المغذية خوارزمياتها باللامساواة، واللاشفافية، في دراسة ملفات المتقدمين للوظيفة، آثار سلبية على نفسيتهم بصفة عامة، حيث قد ينحرفون ويلجؤون إلى استعمال أساليب غير مشروعة للحصول على المال، وعلى تطوير قدراتهم على وجه الخصوص، ذلك أنه في الغالب لا يتم إعطاؤهم سبب الرفض، وهو ما فيه منع من تحسين وتطوير قدراتهم واكتساب مهارات جديدة تساعدهم مستقبلاً في إيجاد وظيفة.

<sup>(1)</sup> كيف يمكن لتطبيقات الذكاء الاصطناعي أن تكون متحيزة مقال منشور على الموقع:

. بتاريخ ٢٠٢١/٩/١٩م، تاريخ الاطلاع ٢٠٢٤/٧/١٩م [/https://masaar.net/ar](https://masaar.net/ar)

<sup>(2)</sup> فوكس نيوز هي قناة إعلامية إخبارية أمريكية تُعرض في الولايات المتحدة والعالم من خلال الباقات الأساسية لشبكات التلفزيون الكبلية الأرضية وهي كذلك فضائية إخبارية تبث عن طريق الأقمار الاصطناعية، وتملكها مجموعة فوكس الترفيهية وهي شركة تابعة لشركة فوكس للقرن 21. موسوعة ويكيبيديا الحرة، منشورة على

الموقع: <https://ar.wikipedia.org/wiki>

<sup>(3)</sup> سوزانا شلينبرج: التحيز الخوارزمي وتصعيد اللامساواة، ترجمة: زياد الحازمي، تحرير: تركي طوهري، مقال <https://atharah.net/how-biased-algorithms-perpetuate-inequality>م منشور على. بتاريخ

2023/10/13م، تاريخ الاطلاع 2020/07/17م

## المبحث الثاني

### آليات الحد من التحيز الخوارزمي

في محاولة للتغلب على مشكلة تحيز الذكاء الاصطناعي نشرت شركة Innodata المتخصصة في هندسة البيانات 5 ممارسات يمكن الاعتماد عليها للحد من تحيز الذكاء الاصطناعي<sup>(١)</sup> والوصول إلى نماذج تعلم آلي أكثر إنصافاً وشمولاً، في سبيل تبسيط آليات الحد من التحيز الخوارزمي يمكن تقسيم هذا المبحث الى قسمين الأول وذلك على النحو الآتي:

### المطلب الاول

#### ضوابط الحد من التحيز الخوارزمي

أضافة قانون الأسواق الرقمية الأوروبي ٢٠٢٢، مجموعة من الضوابط الواجب اتباعها من اجل الحد من آثار التحيز الخوارزمي سيتم تطبيقها في جميع انحاء الاتحاد الأوروبي لتكون مجازا عالميا محتملاً في إدارة المنصات على الإنترنت. يهدف قانون الخدمات الرقمية<sup>(٢)</sup> (DSA) كذلك إلى إنهاء حقبة كانت تضطلع فيها شركات التكنولوجيا منفردة بوضع سياساتها الخاصة حول إدارة وتحميل المحتوى دون رقابة خارجية فعالة عليها وإصدار تقارير شفافية تكون فيها هي الخصم والحكم في أن واحد، ليصبح هناك آليات جديدة من شأنها إجبار المنصات على الإنترنت أن تكون أكثر شفافية بشأن إدارتها للمحتوى الرقمي، وكيفية عمل أنظمتها الخوارزمية، ومحاسبتها على المخاطر المجتمعية الناجمة عن استخدام خدماتها من خلال نظام رقابي مستقل محايد تخضع له مرة واحدة على الأقل سنوياً وفقاً لما نصت عليه المادة ٣٧ من القانون.

كذلك الزم القانون كل مؤسسة من منصات أو محركات البحث الضخمة على الإنترنت بتعيين موظف تحت مسمى (منسق الخدمات الرقمية - Digital Services Coordinator). والزمهم بأن يوفر له إمكانية الوصول إلى البيانات اللازمة والأنظمة الخوارزمية لرصد وتقديم مدى الالتزام بقانون الخدمات الرقمية الأوروبي الجديد وفق ما أشارت إليه المادة ٩٦ من مواد الإصدار، والفترة الأولى من المادة ٣٧ من قانون الأسواق الرقمية الأوروبي<sup>(٣)</sup> هو قانون جديد ثم اعتماده من قبل الاتحاد الأوروبي في يوليو ٢٠٢٢ بهدف إلى معالجة الممارسات الاحتكارية للمنصات الرقمية الكبيرة بهدف القانون إلى خلق بيئة أكثر إنصافاً للمنافسة في السوق الرقمية، وحماية حقوق المستخدمين بعد قانون الأسواق الرقمية

(١) محمد فحي: كيف يمكن التغلب على تحيز الذكاء الاصطناعي..... تقنيات وأدوات، مرجع سابق.

(٢) اللائحة (الاتحاد الأوروبي) ١٩٢٥/٢٠٢٢ الصادرة عن البرلمان الأوروبي والمجلس بتاريخ ١٤ سبتمبر ٢٠٢٢ بشأن الأسواق القابلة للتنافس والعدالة في القطاع الرقمي وتعديل التوجيهين (الاتحاد الأوروبي) ١٩٣٧/٢٠١٩ و(الاتحاد الأوروبي) ١٨٢٨/٢٠٢٠ (قانون الأسواق الرقمية)

(٣) المادة ٣٧: التدقيق المستقل: ١. يخضع مزودو المنصات الإلكترونية الكبيرة جداً ومحركات البحث الإلكترونية الكبيرة جداً، على نفقتهم الخاصة، ومرة واحدة على الأقل سنوياً، لعمليات تدقيق مستقلة لتقييم الالتزام بما يلي

الأوروبي خطوة مهمة في الجهود الرامية إلى لتنظيم الانترنت بشكل أفضل. يسלט الضوء على أهمية المنافسة العائلة في السوق الرقمية، ويوفر إطاراً جديداً لمكافحة الممارسات الاحتكارية للمنصات الرقمية الكبيرة.

حيث شكل هذا القانون تحولا جذرياً في تنظيم الانترنت في أوروبا، والذي يهدف إلى جعل الإنترنت مكاناً أكثر أماناً ومسؤولية للمستخدمين، من خلال فرض قواعد جديدة وبالأخص على عمالقة التكنولوجيا في العالم والتي يستخدمها ٤٥ مليون مستخدم نشط شهرياً أو أكثر في الاتحاد الأوروبي وفقاً لما نصت على الفقرة الأولى من المادة ٣٣ من القانون.

يرى البعض أن قانون الخدمات الرقمية (DSA) عبارة عن مجموعة جديدة من القواعد التي تهدف إلى إجبار منصات الإنترنت الرئيسية مثل Facebook و YouTube و TikTok و Twitter وغيرها على بذل المزيد من الجهد المعالجة انتشار المحتوى غير القانوني والمخاطر المجتمعية الأخرى على خدماتها في الاتحاد الأوروبي لا سيما مخاطر خوارزمياتها، وإلا ستعاقب بغرامات تقدر بمليارات اليوروهات. بالإضافة إلى قانون الأسواق الرقمية الأوروبي ٢٠٢٢، يحددان مجموعة واحدة من القواعد سيتم تطبيقها في جميع أنحاء الاتحاد الأوروبي لتكون مجازاً عالمياً محتملاً في إدارة المنصات على الإنترنت.

يهدف قانون الخدمات الرقمية (DSA) كذلك إلى إنهاء حقبة كانت تضطلع فيها شركات التكنولوجيا منفردة بوضع سياساتها الخاصة حول إدارة وتحميل المحتوى دون رقابة خارجية فعالة عليها وإصدار تقارير شفافية تكون فيها هي الخصم والحكم في أن واحد، ليصبح هناك آليات جديدة من شأنها إجبار المنصات على الإنترنت أن تكون أكثر شفافية بشأن إدارتها للمحتوى الرقمي، وكيفية عمل أنظمتها الخوارزمية، ومحاسبتها على المخاطر المجتمعية الناجمة عن استخدام خدماتها من خلال نظام رقابي مستقل محايد تخضع له مرة واحدة على الأقل سنوياً وفقاً لما نصت عليه المادة ٣٧ من القانون<sup>(١)</sup>.

كذلك الزم القانون كل مؤسسة من منصات أو محركات البحث الضخمة على الإنترنت بتعيين موظف تحت مسمى (منسق الخدمات الرقمية – Digital Services Coordinator) والزمهم بأن يوفر له إمكانية الوصول إلى البيانات اللازمة والأنظمة الخوارزمية لرصد وتقديم مدى الالتزام بقانون الخدمات الرقمية الأوروبي الجديد وفق ما أشارت إليه المادة ٩٦ من مواد الإصدار، والفترة الأولى من

---

(1) Providers of very large online platforms and of very large online search ; Independent audit Article 37 )2(engines shall be subject, at their own expense and at least once a year, to independent audits to assess compliance with the following:

(a) the obligations set out in Chapter III;

(b) Any commitments undertaken pursuant to the codes of conduct referred to in Articles 45 and 46 and the crisis protocols referred to in Article 48.....”

المادة ٤٠ من قانون الأسواق الرقمية الأوروبي<sup>(١)</sup> هو قانون جديد تم اعتماده من قبل الاتحاد الأوروبي في يوليو ٢٠٢٢ بهدف إلى معالجة الممارسات الاحتكارية للمنصات الرقمية الكبيرة بهدف القانون إلى خلق بيئة أكثر إنصافاً للمنافسة في السوق الرقمية، وحماية حقوق المستخدمين بالإضافة إلى قانون الأسواق الرقمية الأوروبي<sup>(٢)</sup> خطوة مهمة في الجهود الرامية إلى لتنظيم الإنترنت بشكل أفضل.

يسلط الضوء على أهمية المنافسة العالمية في السوق الرقمية، ويوفر إطاراً قانونياً جديداً لمكافحة الممارسات الاحتكارية للمنصات الرقمية الكبيرة جدير بالذكر أنه بدأ تفعيل القانون الأوروبي الجديد في ٢٥ أغسطس ٢٠٢٣ من خلال تحديد ١٧ منصة إلكترونية ضخمة (VLOSE)<sup>(٣)</sup>، وعدد ٢ من محركات البحث الضخمة ليكونوا ملتزمين في غضون أربعة أشهر<sup>(٤)</sup>، بمجموعة كاملة من الالتزامات

---

(١) يشير الاختصار VLOPs إلى Very Large Online Platforms ، وتعني منصات ضخمة جدا عبر الإنترنت، وحددها المشرع الأوروبي في الآتي:

: Alibaba AliExpress – Amazon Store – Apple AppStore – Booking.com – Facebook – Google Play –Google Maps – Google Shopping – Instagram – LinkedIn – Pinterest – Snapchat – TikTok – Twitter –Wikipedia – YouTube – Zalando

(٢) يهدف قانون الأسواق الرقمية الأوروبي هو قانون جديد تم اعتماده من قبل الاتحاد الأوروبي في يوليو ٢٠٢٢ للممارسات الاحتكارية للمنصات الرقمية الكبيرة. يهدف القانون إلى خلق بيئة أكثر إنصافاً للمنافسة في السوق الرقمية، وحماية حقوق المستخدمين. يُعد قانون الأسواق الرقمية الأوروبي خطوة مهمة في الجهود الرامية إلى تنظيم الإنترنت بشكل أفضل. يسلط الضوء على أهمية المنافسة العادلة

في السوق الرقمية، ويوفر إطاراً قانونياً جديداً لمكافحة الممارسات الاحتكارية للمنصات الرقمية الكبيرة. الاطلاع على القانون

Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act)

(٣) وتعني محركات بحث منسجمة جداً، Very Large Online Search Engines، يشير الاختصار (VLOSE) وتعني منصات ضخمة جدا عبر الإنترنت، وحددها المشرع الأوروبي مثل:

Very Large Online Platforms (VLOP) يشير الاختصار )Google Play – Facebook – Booking.com Apple AppStore Amazon Store – Alibaba AliExpress في التي Twitter TaTok Snapchat Pinterest Unkedin Instagram Google Shopping Google Maps Zalando YouTube– Wikipedia

(٤) By 4 months after notification of the designated decisions, the designated platforms and search engines need to adapt their systems, resources, and processes for compliance, set up an independent system of compliance and carry out, and report to the Commission, their first annual risk assessment. Digital Services Act:

الجديدة بموجب هذا القانون أهمها تقييم المخاطر واتخاذ ما يلزم من تدابير للحد والتخفيف منها، على أن يعمم القانون على كافة المنصات الأخرى عبر الضخمة في جميع أنحاء الاتحاد الأوروبي بحلول فبراير ٢٠٢٤ (١)

حدد القانون ٤ فئات مختلفة من المخاطر النظامية systemic risks " التي تسأل عنها المنصات ومحركات البحث الضخمة على الإنترنت بموجب المواد، ٨٣، ٨٢، ٨١، ٨٠، من مواد الإصدار (١)، وهم؛

---

Commission designates first set of Very Large Online Platforms and Search Engines, Accessed 9 – 8–2023, On this website;

[https://ec.europa.eu/commission/presscorner/detail/en/IP\\_23\\_241](https://ec.europa.eu/commission/presscorner/detail/en/IP_23_241)

(١) ولعل أهم هذه الالتزامات الآتي:

- ١- المزيد من تمكين المستخدمين؛ لاسيما تمكين المستخدمين من الإبلاغ عن المحتوى غير القانوني بطرق أكثر بساطة ويسر، حظر الإعلانات بناءً على البيانات الحساسة للمستخدم (مثل الأصل العرقي أو الآراء السياسية أو التوجه الجنسي) ، إبلاغ المستخدمين بمن يقوم بترويج الإعلانات، تقديم ملخص سهل الفهم بلغة واضحة لشروطها وأحكامها، بلغات الدول الأعضاء التي تعمل فيها.
- ٢- حماية إضافية للقاصرين؛ كضمان مستوى عالٍ من الخصوصية والأمن والسلامة للقاصرين، وتقديم تقرير عن المخاطر الخاصة بالقاصرين وإعادة تصميم خدماتها بما يخفف من هذه المخاطر.
- ٣- الحد من التضليل المعلوماتي؛ يتعين على المنصات ومحركات البحث اتخاذ تدابير اللازمة لمعالجة المخاطر المرتبطة بنشر المحتوى غير القانوني عبر الإنترنت والآثار السلبية على حرية التعبير والمعلومات، كذلك أن تجد آلية تتيح للمستخدمين الإبلاغ عن المحتوى غير القانوني والتصرف بناءً على الإشعارات بسرعة.
- ٤- المزيد من الشفافية والمساءلة؛ على المنصات التأكد من تقييم المخاطر الخاصة بها وامتثالها لجميع التزامات التي فرضها القانون الأوروبي الجديد، إتاحة الوصول إلى البيانات للجمهور وللباحثين، نشر تقارير الشفافية حول قرارات الإشراف على المحتوى وإدارة المخاطر.

=For online platforms, they must publish their number of active users by February 2023. If the platform or a search engine has more than 45 million users (10% of the population in Europe), the Commission will designate the service as a very large online platform or a very large online search engine. These services will have 4 months to comply with the obligations of the DSA, which includes carrying out and providing the Commission with their first annual risk assessment. EU Member States will have to appoint Digital Services Coordinators by 17 February 2024, when also platforms with less than 45 million active users have to comply with all the DSA rules. The Digital Services Act package, Shaping Europe's digital future, Accessed 8–9–2023, On this website; <https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package>

الفئة الأولى: المخاطر المرتبطة بنشر محتوى غير قانوني <sup>(٢)</sup> الفئة الثانية المخاطر المرتبطة بالتأثير الفعلي أو المتوقع للخدمة على ممارسة الحقوق الأساسية كالكرامة الإنسانية، وحرية التعبير والمعلومات،

---

<sup>(١)</sup> حيث تكون القانون من 156 مادة من مواد الإصدار، و 93 مادة أخرى تمثل متن القانون وزعت على 5 فصول لكل فصا أقسامه ومواده المتعلقة به.

<sup>(٢)</sup> وهو ما نصت عليه المادة 80 من مواد الإصدار؛ مثل نشر مواد الاعتداء الجنسي على الأطفال أو خطاب الكراهية غير القانوني أو أنواع أخرى من إساءة استخدام خدماتهم في جرائم جنائية، وممارسة أنشطة غير قانونية، مثل بيع المنتجات أو الخدمات المحظورة بموجب القانون الاتحادي أو القانون الوطني، بما في ذلك المنتجات الخطرة أو المقلدة أو الحيوانات التي يتم الاتجار بها بشكل غير قانوني.

" 80 ) Four categories of systemic risks should be assessed in-depth by the providers of very large online platforms and of very large online search engines. A first category concerns the risks associated with the dissemination of illegal content, such as the dissemination of child sexual abuse material or illegal hate speech or other types of misuse of their services for criminal offences, and the conduct of illegal activities .....",

(81) " A second category concerns the actual or foreseeable impact of the service on the exercise of fundamental rights, as protected by the Charter, including but not limited to human dignity, freedom of expression and of information, including media freedom and pluralism, the right to private life, data protection, the right to non-discrimination, the rights of the child and consumer protection. Such risks may arise, for example, in relation to the design of the algorithmic systems used by the very large online platform or by the very large online search engine ....".

(82)(4) " Processes, A third category of risks concerns the actual or foreseeable negative effects on democratic civic discourse and electoral processes, as well as public security.. ..."

(83)(5)" A fourth category of risks stems from similar concerns relating to the design, functioning or use, including through manipulation, of very large online platforms and of very large online search engines with an actual or foreseeable negative effect on the =protection of public health, minors and serious negative consequences to a person's physical and mental well-being, or on gender-based violence. Such risks may also stem from coordinated disinformation campaigns related to public health, or from online interface design that may stimulate behavioral addictions of recipients of the service..

بما في ذلك حرية وسائل الإعلام والتعددية، والحق في الخصوصية والحياة، وحماية البيانات، والحق في عدم التمييز وحقوق الطفل، وحماية المستهلك، والتي قد تنشأ من تصميم الأنظمة الخوارزمية الفئة الثالثة: المخاطر المرتبطة بالآثار السلبية الفعلية أو المتوقعة على الديمقراطية والعمليات الانتخابية والأمن العام). أما الفئة الرابعة والأخيرة المخاطر المرتبطة بالتأثير السلبي على الصحة العامة والقاصرين والأطفال<sup>(١)</sup>.

أوجبت المادة ٨٤ من مواد الإصدار أنه على مقدمي المنصات ومحركات البحث الضخمة على الإنترنت عند تقييم هذه المخاطر النظامية systemic risks " التركيز على ضبط ومعالجة وإصلاح الأنظمة أو العناصر الأخرى التي قد تساهم في تضخم هذه المخاطر<sup>(١)</sup>، بما في ذلك جميع الأنظمة الخوارزمية التي قد تكون ذات صلة<sup>(٢)</sup>. كما أوجبت المادة ٨٨ على مقدمي المنصات ومحركات البحث الضخمة على الإنترنت أن يبذلوا قصارى جهدهم في التدابير التي يتخذونها لاختبار أنظمتهم الخوارزمية، وتكييفها عند الضرورة<sup>(٣)</sup>.

وفي حال مخالفة أي التزام من الالتزامات التي شملها القانون لا سيما الالتزام بتحديد وتصحيح أي تحيزات قد تكون موجودة في خوارزمياتها، توقع غرامة تقدر ب ٦% من إجمالي المبيعات السنوية للمنصات أو محركات البحث الضخمة على مستوى العالم، ويتم احتساب هذه النسبة من مبيعات السنة

---

<sup>(١)</sup> حيث نصت المادة 154 من مواد الإصدار على أن " نظراً لنطاق وتأثير المخاطر المجتمعية التي قد تسببها منصات ومحركات الإنترنت الضخمة فإن الحاجة إلى معالجة تلك المخاطر على سبيل الأولوية والقدرة على اتخاذ التدابير اللازمة"

**Given the scope and impact of societal risks that may be caused by very large online platforms and very large online search engines, the need to address those risks as a matter of priority and the capacity to take the necessary measures, it is justified to limit the period after which this Regulation starts to apply to the providers of those services.**

**(2)When assessing such systemic risks, providers of very large online platforms and of very large on search engines should focus on the systems or other elements that may contribute to the risks, including all the algorithmic systems that may be relevant, in particular their recommender systems and advertising systems, paying attention to the related data collection and use practices.**

**(3)Providers of very large online platforms and of very large online search engines should also be diligent in the measures they take to test and, where necessary .".**

المالية السابقة<sup>(١)</sup>، وفقاً لما نصت عليه المادة ٧٤ الموسومة بـ (الغرامات)<sup>(١)</sup> يفرض إذن قانون الخدمات الرقمية الأوروبي الجديد (DSA) غرامات على الشركات الكبرى في حالة التحيز الخوارزمي. ويحدد القانون معايير صارمة لمنع هذا النوع من التمييز العنصري المستحدث. وإذا فشلت الشركة في الوفاء بهذه المتطلبات، فقد تواجه غرامات تصل إلى ٦% من عائداتها العالمية. من ثم يهدف القانون الأوروبي الجديد DSA إلى حماية المستخدمين من التحيز الخوارزمي، ويوفر إطاراً تنظيمياً شاملاً لمعالجته.

يتطلب من الشركات الكبرى اتخاذ إجراءات وقائية لمنع التحيز، ويمنح المنظمين سلطة اتخاذ إجراءات ضد الشركات التي لا تلتزم بالمعايير التي يمكن تلخيصها في الآتي:

١- تحديد وتقييم المخاطر يجب على الشركات الكبرى تحديد وتقييم المخاطر المحتملة للتحيز في خوارزمياتها. ويجب أن تأخذ الشركات في الاعتبار مجموعة متنوعة من العوامل، بما في ذلك معايير الخوارزمية، والبيانات المستخدمة في تدريب الخوارزمية، والنتائج المحتملة للتحيز<sup>(٢)</sup>.

---

<sup>(١)</sup> كما أقرت المادة ذاتها غرامة أخرى لا تتجاوز 1% من الإجمالي الدخل السنوي أو حجم الأعمال العالمي في السنة المالية السابقة في حالات محددة سواء كانت عن قصد أو عن خطأ وهي

- ١- تقديم معلومات غير صحيحة أو غير كاملة أو مضللة بالمخالفة لأحكام المادة 67 رفض الخضوع للتفتيش عملاً بالمادة 69 ؛
- ٢- عدم الرد على طلب المعلومات بقى ا رر خلال الفترة المحددة رفض الخضوع للتفتيش عملاً بالمادة 69 ؛
- ٣- عدم تصحيح المعلومات غير الصحيحة أو غير الكاملة أو المضللة التي قدمها أحد الموظفين خلال الفترة التي حددتها اللجنة، أو رفض تقديم معلومات كاملة؛
- ٤- رفض الخضوع للتفتيش عملاً بالمادة 69 ؛
- ٥- عدم الامتثال للتدابير التي اعتمدها اللجنة عملاً بالمادة 72 ؛
- ٦- عدم الالتزام بشروط الوصول إلى ملف اللجنة عملاً بالمادة 79 الفقرة 4

<sup>(٢)</sup> ومن أجل مراقبة وتقييم التزام منصات ومحركات البحث الضخمة على الإنترنت بعدم التحيز الخوارزمي، قد يطلب منسق الخدمات الرقمية في المؤسسة أو اللجنة الوصول إلى بيانات محددة أو الإبلاغ عنها، بما في ذلك البيانات المتعلقة بالخوارزميات، كالبيانات اللازمة لتقييم المخاطر والأضرار المحتملة الناجمة عن الأنظمة الأساسية، وبيانات حول دقة وعمل واختبار الأنظمة الخوارزمية للإشراف على المحتوى. أو أنظمة التوصية أو أنظمة الإعلان، كذلك بيانات تدريب الخوارزميات، وغيرها. وفقاً لما نصت عليه المادة 96 من مواد الإصدار؛

(96) In order to appropriately monitor and assess the compliance of very large online platforms and of very large online search engines with the obligations laid down by this Regulation, the Digital Services Coordinator of establishment or the Commission may require access to or reporting of specific data, including data related to algorithms. Such a requirement may include, for example, the data necessary to

٢- اتخاذ تدابير وقائية يجب على الشركات الكبرى اتخاذ تدابير وقائية لمنع التحيز في خوارزمياتها. يمكن أن تشمل هذه التدابير ما يلي: استخدام بيانات تدريب متنوعة وتمثيلية، استخدام الخوارزميات المستقلة عن العوامل الشخصية مراقبة أداء الخوارزميات وتحديد التحيزات المحتملة.

٣- الإبلاغ عن التحيزات يجب على الشركات الكبرى الإبلاغ عن أي تحيزات يتم تحديدها في خوارزمياتها إلى المفوضية الأوروبية. ويجب أن يتضمن الإبلاغ وصفا للتحيز، وتحليلاً للأسباب المحتملة، وخطاً لتصحيح التحيز.

في حالة الإخلال بأحد المعايير السابقة، تفرض على الشركات الكبرى والمسؤولة عن المنصات ومحركات البحث الضخمة على الإنترنت غرامة تصل إلى ٦% من عائداتها العالمية تحتسب على أساس السنة المالية السابقة.

وحتى يكون قانون الخدمات الرقمية الأوروبي أكثر نجاعة وفعالية في الحد من التحيز الخوارزمي انشأت المفوضية الأوروبية المركز الأوروبي للشفافية الخوارزمية ( The European Centre for Algorithmic Transparency (ECAT) ليضطلع بدور فني مهم في الإشراف والرقابة على دقة وحيادية خوارزميات المنصات ومحركات البحث الضخمة على الإنترنت، وبدأ العمل به رسمياً في إبريل ٢٠٢٣.

ويرى الباحث ضرورة وضع اليات محدد للحد من التحيز الخوارزمي من ضمن هذا الضوابط:

أولاً: اختيار مجموعة البيانات:

في حين أن التخفيف من حدة تحيز الذكاء الاصطناعي والتعلم الآلي يمكن أن يمثل تحدياً، إلا أن هناك تقنيات وقائية يمكن أن تساعد في حل هذه المشكلة. التحدي الأكبر في تحديد التحيز يتمثل في

---

assess the risks and possible harms brought about by the very large online platform's or the very large online search engine's systems, data on the accuracy, functioning and testing of algorithmic systems for content moderation, recommender systems or advertising systems, including, where appropriate, training data and algorithms, or data on processes and outputs of content moderation or of internal complaint-handling systems within the meaning of this Regulation. Such data access requests should not include requests to produce specific information about individual recipients of the service for the purpose of determining compliance of such recipients with other applicable Union or national law. Investigations by researchers on the evolution and severity of online systemic risks are particularly important for bridging information asymmetries and establishing a resilient system of risk mitigation, informing providers of online platforms, providers of online search engines, Digital Services Coordinators, other competent authorities, the Commission and the public.

فهم كيف تقوم بعض خوارزميات التعلم الآلي بتعميم البيانات التي تم تدريبها عليها. لذلك من المهم أن تتسم البيانات المستخدمة لتدريب النموذج بالشمول والحياد في ذات الوقت.

### ثانياً: تنوع الفريق:

بناء فريق متنوع يساهم إلى حد كبير في القضاء على التحيز، يمكن أن يكون لتنوع الفريق تأثير إيجابي على نماذج التعلم الآلي من خلال إنتاج مجموعات بيانات مُمتلئة ومتوازنة (Representative Datasets). كما أنه يساعد في التخفيف من التحيز الضار في بنية مجموعات البيانات وكيفية تطبيق التصنيفات على تلك البيانات.

### ثالثاً: تقليل الإقصاء :

يعد اختيار الميزة (Feature selection) أمراً أساسياً للمساعدة في تقليل الإقصاء في الذكاء الاصطناعي، وهي عملية تقوم على تقليل عدد المتغيرات المُدخل إلى نماذج الذكاء الاصطناعي بهدف تحسين أداء التنبؤ بالنتائج. تستبعد هذه الخطوة عناصر البيانات التي لا تحتوي على تباين كافٍ للتأثير على النتائج.

### رابعاً: الخوارزميات وحدها ليست كافية:

هناك طريقة أخرى للبدء في حل مشكلة التحيز في الذكاء الاصطناعي، وهي عدم الاعتماد فقط على الخوارزميات، بل إبقاء الأفراد القائمين على تطوير الذكاء الاصطناعي على اطلاع بكل ما يتعلق بالنظام ليتمكنوا من التعرف بفعالية على أنماط التحيز غير المقصود. يمكن أن تعمل هذه الخطوة على تقليل العيوب الموجودة في النظام، مما يخلق نموذج تعلم آلي أكثر حيادية.

يجب على المنظمات أيضاً وضع مبادئ توجيهية وإجراءات تحدد التحيز المحتمل في مجموعات البيانات وتخفف من حدته. يمكن أن يساعد توثيق حالات التحيز عند حدوثها وتحديد كيفية العثور عليها والحديث بشأنها، على ضمان عدم تكرارها.

### خامساً: بيانات ممثلة:

يجب على المؤسسات فهم كيف تبدو البيانات الممثلة قبل جمع البيانات التي سيتم تدريب نموذج التعلم الآلي عليها. كما يجب أن يحتوي جوهر وخصائص البيانات المستخدمة أقل قدر من التحيز. وإلى جانب تحديد التحيز المحتمل في مجموعات البيانات، يجب على المؤسسات أيضاً توثيق أساليبها في اختيار البيانات وتنقيتها، للقضاء جذرياً على أسباب التحيز.

**المطلب الثاني**  
**الآليات الخاصة بالحد من التحيز الخوارزمي**  
**في مجال التوظيف وحقوق الإنسان**  
**الفرع الأول**  
**وضع اليات قانونية وتنظيمية**

نظرا للانتهاكات الواقعة على حق الفرد في الحصول حق التوظيف باعتباره حق من حقوق الإنسان، فقد ينتهك ذلك الحق من قبل أصحاب الشركات والقرارات شهد مجلس الشيوخ الأمريكي سنة 2019م تقديم مشروع قانون المساواة الخوارزمية من طرف النواب كوري بوكر، ورون ويدن، ويفت ديفت كلارك، والذي يلزم الشركات بتقييم وإصلاح الخوارزميات الحاسوبية المعيبة، التي تؤدي إلى قرارات غير دقيقة أو غير عادلة أو تمييزية، وإعطاء صورة أكثر شفافية وعدالة ومساواة لاستخدام الخوارزميات، حيث قالت كلارك: "الخوارزميات لا يجب أن تكون معفاة من قوانيننا التي تحارب التمييز"<sup>(1)</sup>.

وفي سبيل الحفاظ على حقوق الإنسان جراء الاستخدام الواسع للذكاء الاصطناعي عموما، وإقامة العدل والمساواة بين الباحثين عن العمل بغض النظر عن عرقهم أو جنسهم أو دينهم، تدخلت منظمة اليونسكو سنة 2021م، وأصدرت توصية خاصة بشأن أخلاقيات الذكاء الاصطناعي<sup>(2)</sup>، التي تعد أول إطار عالمي يعالج موضوع الاستخدام الأخلاقي للذكاء الاصطناعي، حيث جاء في العنصر الموسوم بالقيم احترام حقوق الإنسان والحريات الأساسية وكرامة الإنسان وحمايتها وتعزيزها في الفقرة 13 منه ما يلي: "تعد كرامة الإنسان، التي لا يجوز المساس بها، والتي ينبغي لكل إنسان أن يتمتع بها، القاعدة التي تقوم عليها منظومة حقوق الإنسان والحريات الأساسية العالمية المترابطة والمتشابكة وغير القابلة للتجزئة والتصرف ولذلك لا بدّ من احترام كرامة الإنسان وحقوق الإنسان وحمايتها وتعزيزها وفقاً لأحكام القانون الدولي، ومنها أحكام القانون الدولي لحقوق الإنسان، طوال دورة حياة نُظَم الذكاء الاصطناعي.

وترتبط كرامة الإنسان بالإقرار بالقيمة الذاتية الملازمة لكل إنسان، التي يتساوى فيها الناس كافة بغض النظر عن العرق أو اللون أو النسب، أو الجنس أو السن، أو اللغة أو الدين، أو الآراء السياسية، أو الأصل القومي أو الإثني أو الاجتماعي، أو الوضع الاقتصادي أو الاجتماعي المرتبط بالمولد، أو الإعاقة، أو أي سبب من أسباب التمييز الأخرى.

وجاء في الفقرة 16 منها كذلك: "ويجب احترام حقوق الإنسان والحريات الأساسية وحمايتها وتعزيزها طوال دورة حياة نُظَم الذكاء الاصطناعي. ويجب على الحكومات، وشركات القطاع الخاص،

---

(1) Booker, Wyden, Clarke Introduce Bill Requiring Companies To Target Bias In Corporate Algorithms, published on <https://www.booker.senate.gov/>, ٢٠١٩/٤/١٠

التوصية الخاصة بأخلاقيات الذكاء الاصطناعي منشورة على موقع منظمة اليونسكو

<sup>2-</sup> [/https://unesdoc.unesco.org](https://unesdoc.unesco.org)

وهيئات المجتمع المدني، والمنظمات الدولية، والأوساط التقنية والأكاديمية، أن تحترم صكوك وأطر حقوق الإنسان في أعمالها المتعلقة بالعمليات المرتبطة بدورة حياة نُظم الذكاء الاصطناعي . ويجب أن توفر الوسائل التكنولوجية الجديدة أدوات جديدة لمناصرة حقوق الإنسان والدفاع عنها وممارستها، ويجب ألا تؤدي إلى انتهاك هذه الحقوق ."

وبهذه النصوص الواردة في متن التوصية، تكون منظمة اليونسكو قد أحسنت فيما فعلت في مبادرتها الأولى للتأكيد على احترام القيم، وحقوق الإنسان إزاء الاستخدام الأخلاقي لأنظمة الذكاء الاصطناعي، التي تشهد تطوراً متلاحقاً يوماً بعد يوم، وعليه يكون حرياً بالدول العمل على ترسيخ حقوق الإنسان في صميم الأطر التنظيمية والتشريعات الخاصة بتطوير واستخدام تكنولوجيات الذكاء الاصطناعي.

## الفرع الثاني

### دمج الإشراف البشري في عملية التوظيف

بما أن أغلب الشركات أصبحت تعتمد على أنظمة الذكاء الاصطناعي في اختيار العمال الراغبين بالعمل لديها، الأمر الذي يؤدي بالضرورة إلى عدم وجود عدل ومساواة في دراسة الملفات، فقد يوظف من هو غير مؤهل للوظيفة المطروحة، وهو ما ينعكس بالسلب على أداء الشركة، ولتجنب الوقوع في هذه المشكلة، وإعطاء كل شخص فرصة لدراسة ملفه استناداً إلى معايير موضوعية بعيدة عن كل تحيز عرقي أو جنس ي أو ديني أو تاريخي، يتعين على الشركات دمج الإشراف والتدخل البشري في عملية التوظيف.

## الفرع الثالث

### تدريب أنظمة التوظيف الآلي

يجب على أرباب العمل القيام بتدريب أنظمة التوظيف الآلي القائمة على أنظمة الذكاء الاصطناعي بمجموعة متنوعة ومتوازنة من البيانات، بحيث تشمل كل أصناف الأفراد باختلاف عرقهم وجنسهم ودينهم، وحالاتهم والاجتماعية، التعليمية، والتاريخية... إلخ، مما يعطي فرصاً وحظوظاً أكبر للباحث عن العمل الحائز لجميع الشروط المطلوبة في الوظيفة.

## الفرع الرابع

### مراقبة وتقييم أنظمة التوظيف الآلي

تعد أنظمة التوظيف الآلي أنظمة مؤتمتة، وبالتالي فهي تحتفظ بجميع المعلومات التي قامت بمعالجتها سابقاً، فإذا سبق لها إصدار قرارات منحازة بناء على معطيات مبنية على فوارق اجتماعية، وعرقية، ودينية، وتاريخية، فإن هذه المعطيات المنحازة تتأصل فيها، وهو ما يستوجب عمل مراقبة دائمة ومستمرة بحثاً عن التحيزات الموجودة وعمل تحديث لها وفق ضوابط أخلاقية.

## المطلب الثالث

### تشريعات الحد من التحيز

تمهيد وتقسيم:

لبيان ما قمت بها التشريعات الأوروبية، سنتناول في هذا المطلب ونقسمه الى فرعين؛ الأول: نخصه في الاتحاد الأوروبي، والثاني: في بعض الدول العربية.

### الفرع الأول

#### تشريعات الحد في الاتحاد الأوروبي

لا بد من الإشارة إلى أنه فيما يتعلق بالآثار المترتبة على استخدام أنظمة الذكاء الاصطناعي لحماية البيانات الشخصية، فإن مجموعة الإشكالات التي تحول دون التأثيرات السلبية لخوارزميات الذكاء الاصطناعي<sup>1</sup> تتضمن جملة من الحقوق، منها: الموافقة الصريحة لأصحاب البيانات على معالجة بياناتهم الشخصية<sup>(1)</sup>، ومبدأ تقليل البيانات ومبدأ تحديد الغرض ومجموعة الحقوق المتعلقة بوقت السماح باتخاذ القرار المؤتمنة.

أولاً: موقف اللائحة العامة لحماية البيانات الأوروبية من العدالة الخوارزمية.

حيث دخلت اللائحة التنظيمية الأوروبية رقم ٩٦٧/ ٢٠١٦ حيز التنفيذ في ٢٥ مايو ٢٠١٨ وأدرجت آلياً بالنظام التشريعي الوطني لكافة دول الاتحاد الأوروبي<sup>(2)</sup>، وأصبحت جزءاً لا يتجزأ من اللائحة التنظيمية الأوروبية، وهي الآن سارية المفعول بالبلدان المعنية وستوفر حماية عليا للمعطيات الشخصية بالمنطقة الأوروبية، ففي حالات المعالجة الخوارزمية مثلاً، ينبغي على مراقب البيانات تنفيذ تدابير مناسبة لحماية حقوق وحرية ومصالح أصحاب البيانات ومصالحهم المشروعة، من خال ضمان حقوقهم في الحصول على تدخل إنساني لدى وحدة التحكم والمعالجة ببرمجية الخوارزميات للتعبير عن حقهم، في الطعن في القرارات المؤتمنة<sup>المادة ٢٢</sup>، الفقرة ٣ من اللائحة العامة لحماية البيانات

<sup>(1)</sup> تجدر الإشارة إلى أن قانون حماية البيانات الشخصية المصري رقم ١٥١ لسنة ٢٠٢٠ (نص في المادة (١) منه على تعريف البيانات الشخصية بأنها: (أي بيانات متعلقة بشخص طبيعي محدد، أو يمكن تحديده بشكل مباشر أو غير مباشر عن طريق الربط بين هذه البيانات وأي بيانات أخرى كالاسم، أو الصوت، أو الصورة أو رقم تعريف، أو محدد للهوية عبر الإنترنت، أو بيانات تحدد الهوية النفسية أو الصحية، أو الاقتصادية أو الثقافية، أو الاجتماعية).

<sup>(2)</sup> Veale M, et,al, 'Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision Making' in p.4. doi: 10/ct4s.Proceedings of the ACM Conference on Human Factors in Computing Systems, CHI 2018 (ACM Press )8102, 7,

الأوروبية<sup>(١)</sup>، وتتضمن اللائحة العامة لحماية البيانات حق الشخص الخاضع لنظام المعالجة الخوارزمية لتلقي بيانات ذات معنى حول المنطق الذي تضمن المعالجة الآلية لبياناته الشخصية<sup>(٢)</sup>؛ حيث تتطلب المواد ٢/١٣، و٢/١٤ ز، و١/١٥ ح من اللائحة العامة لحماية البيانات مراقبي البيانات تزويد صاحب البيانات بمعلومات حول وجود أتمته للقرارات، المشار إليها في المادة ١/٢٢، و٤/٢٢، والتي تنص على أنه: "يحق لصاحب البيانات ألا يوافق على قرار يعتمد فقط على المعالجة الآلية، بما في ذلك التتميط، والذي ينتج عنه آثاراً قانونية تتعلق بها وتؤثر عليه بشكل جسيم ...

بالإضافة إلى أهمية هذه المعالجة وعواقبها المتصورة على صاحب البيانات، فقد جاء في المادتين ١٣ و ١٤ على مراقبي البيانات إخطار أصحاب البيانات الشخصية بالمعالجة الخوارزمية لبياناتهم، وورد في المادة ١٥ على أنها تعطي الحق لهم في الوصول إلى المعلومات الشخصية مدة المعالجة الخوارزمية للبيانات<sup>(٣)</sup>، والحق في عدم الخضوع لاتخاذ قرارات خوارزمية مع اتخاذ العديد من الإجراءات الوقائية والقيود للحالات المحدودة التي يُسمح فيها باتخاذ القرارات المؤتمتة.

وأطلقت المفوضية الأوروبية لكفاءة العدالة "CEPEJ" التابعة لمجلس أوروبا "الميثاق الأخلاقي الأوروبي بشأن استخدام الذكاء الاصطناعي في النظم القضائية وبيئتها"<sup>(٤)</sup>؛ حيث ناقش الميثاق الأوروبي لأخلاقيات استخدام الذكاء الاصطناعي في الأنظمة القضائية لعام ٢٠١٨ هذه الفرضية، مبيناً الخصوصية للأنظمة القانونية في الدول الأوروبية، وخطورة هذه العملية على الصورة الذهنية التي يمكن أن تتكون حول القاضي، ببعديها السلبي والإيجابي، وأثر ذلك على شخص القاضي والغير، والخشية من أن يكون ذلك نوعاً من التشهير غير المباشر على القاضي من خلال الحكم المسبق عليه، وجاء في الميثاق النص على المبادئ الخمسة للميثاق الأخلاقي على استخدام الذكاء الاصطناعي في النظم القضائية وبيئتها، وهي: مبدأ احترام الحقوق الأساسية: من خلال ضمان أن تصميم وتنفيذ أدوات الذكاء الاصطناعي متوافقة مع الحقوق الأساسية للإنسان، ومبدأ عدم التمييز: عبر منع تطوير برامج متحيزة؛

---

(1) Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive

(2) ibid.

(3) Andrew D Selbst & Julia Powles. Meaningful information and the right to explanation, *International Data Privacy Law*, Volume 7, Issue 4, November 2017, Pages 233–242, <https://doi.org/10.1093/idpl/ix022>. (last visited ٢٠٢٠/٥/١).

(4) Reports are available at:

[http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap\\_47\\_2017\\_verkkoju-](http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_verkkoju-)

أي تميز بين الأفراد، وكذلك، مبدأ الجودة والأمن: فيما يتعلق بتصميم الخوارزميات لإصدار القرارات والبيانات القضائية، من خلال استخدام مصادر معتمدة وبيانات آمنة في البيئة التكنولوجية، إضافة إلى مبدأ الشفافية والحياد والنزاهة من خلال وسائل معالجة البيانات بحيث يمكن الوصول إليها وفهماها؛ أي إمكانية مراجعتها من قبل الغير.

وفقاً للمادة ١١ من توجيه إنفاذ القانون الأوروبي، تُحظر القرارات المؤتمتة التي ينتج عنها آثار قانونية سلبية تتعلق بموضوع البيانات أو يكون لها تأثير على حقوق الأفراد، ما لم يأذن بها قانون الاتحاد الأوروبي أو أحد الدول الأعضاء فيه، والذي يجب عليه أيضاً التقيد بالضمانات المناسبة لحقوق وحرية صاحب البيانات، تماشياً مع أحكام توجيه إنفاذ القانون، ولا يمكن أن تكون القرارات القضائية الصادرة بالكامل بواسطة أداة خوارزمية قانونية. وتبنى البرلمان الأوروبي قراراً ينص على أنه: "لا يجوز تبني الخوارزميات في أنظمة صنع القرار بدون تقييم الأثر المسبق للخوارزميات (١)..." فهي لا تنفذ الحكم الوارد في المادة ٢٢/٢، حتى الآن هذا هو الحال في معظم بلدان الاتحاد الأوروبي، على سبيل المثال إيطاليا<sup>(٢)</sup>، رومانيا<sup>(٣)</sup>، السويد<sup>(٤)</sup>، الدنمارك<sup>(٥)</sup>، بولندا<sup>(٦)</sup>، فنلندا<sup>(٧)</sup>، البرتغال<sup>(٨)</sup>، إسبانيا<sup>(٩)</sup>.

---

(1) Deloitte and Reform, the State of the State 2018–19 (2019)

<https://www2.deloitte.com/content/campaigns/uk/thestate-of-the-state/the-state-of-the-state-of-the-state-of-the-state.htm>. –

(2) [http://www.gazzettaufficiale.it/atto/serie\\_generale/Decreto Legislativo 10 Agosto 2018,n.101,caricaDettaglioAtto/originario?atto.dataPubblicazi oneGazzetta=20180904&atto.codiceRedazionale=18G00129&elenco30giorni=tru](http://www.gazzettaufficiale.it/atto/serie_generale/Decreto_Legislativo_10_Agosto_2018,n.101,caricaDettaglioAtto/originario?atto.dataPubblicazi oneGazzetta=20180904&atto.codiceRedazionale=18G00129&elenco30giorni=tru) –

(3) Lege 190/2018 privind nasuri de punere in aplicare a Regulamentului (UE) 2016/679 al Parlamentului European si al Consiliului din 27 aprilie 2016  
<https://www.senat.ro/legis/PDF/2018/18L294FP.pdf>.

(4) [http://www.riksdagen.se/sv/dokumentlagar/dokument/svenskforfattningssamling/for ordning-2018219-med-kompletterande\\_sfs-2018-21](http://www.riksdagen.se/sv/dokumentlagar/dokument/svenskforfattningssamling/for ordning-2018219-med-kompletterande_sfs-2018-21)

(5) <https://www.retsinformation.dk/Forms/R0710.aspx?id=201319>.

(6) <http://prawo.sejm.gov.pl/isap.nsf/download.xsp/WDU20180001000/O/D20181000.pdf>.

(7) <https://likumi.lv/ta/id/300099-fizisko-personu-datu-apstrades-likums>.

(8) <https://likumi.lv/ta/id/300099-fizisko-personu-datu-apstrades-likums>.

(9) [https://iapp.org/media/pdf/resource\\_center/Spanish\\_data-protection-law.pdf](https://iapp.org/media/pdf/resource_center/Spanish_data-protection-law.pdf).

**ثانياً: نهج محايد:** حيث نفذت الدولة العضو المادة ٢/٢٢ من اللائحة العامة لحماية البيانات، بيد أنها لا تتبنى إجراءً مناسباً محددًا لحماية حقوق وحرية صاحب البيانات والمصالح المشروعة. ومثال ذلك: حالة ألمانيا، وجزئياً، النمسا وبلجيكا.

**ثالثاً: نهج اجرائي:** نجد بعض الدول الأوروبية توفر ضمانات محددة بموجب المادة ٢/٢٢ من اللائحة والتي تستند بشكل أساسي إلى وصف الإجراءات التي يجب أن يتخذها الذكاء الاصطناعي في معالجة البيانات الشخصية عند اتخاذ قرارات خوارزمية على الأفراد "مثل حق الإخطار، المراجعة... إلخ" أو بعض أشكال تقييم وتأثير برنامج الخوارزمية، ومثال ذلك، حالة المملكة المتحدة وإيرلندا، وجزئياً، سلوفينيا.

**رابعاً: نهج استباقي:** تقترح بعض الدول الأعضاء ضمانات جديدة وأكثر تحديداً بموجب المادة ٢/٢٢، ومنها: الحق في معرفة آليات عمل الخوارزميات، ومثالها حالة فرنسا والمجر.

## الفرع الثاني تشريعات الحد من التحيز في بعض الدول العربية

في إطار حماية حقوق المتهمون الذين تصدر في حقهم قرارات وأحكام قضائية مبنية على معالجة تلك الخوارزميات المصممة بواسطة الذكاء الاصطناعي، نجد أن المشرع المغربي - قد منح مشروع القانون " المغربي " الصادر في عام ٢٠٠٩" سلطة واسعة بتنفيذ القانون والمتعلق " بحماية الأشخاص تجاه معالجة المعطيات ذات الطابع الشخصي الى إمكانية رفض القرارات القضائية الخوارزمية من ناحية، ومعرفة المنطق الذي يحكم كل معالجة آلية للمعطيات الشخصية التي تخصه من جهة أخرى<sup>(١)</sup>.

وحيث تنص المادة ١١ من القانون المغربي - أنف الذكر - الفرد المعني بالبيانات الشخصية رفض الأحكام المؤتممة التي يترتب عنها نتائج قانونية أو من شأنها أن تؤثر تأثيراً جسيماً في الشخص المعني بها؛ إذ تنص المادة - أنفة الذكر - على أنه: " لا يمكن للأحكام القضائية أن تتضمن تقييماً لسلوك شخص من الأشخاص أن يكون مبنياً على معالجة آلية لمعطيات ذات طابع شخصي يكون الغرض منها تقييم بعض جوانب شخصيته. كذلك، لا يمكن لأي قرار آخر تنشأ عنه آثار قانونية تجاه شخص من الأشخاص أن يتخذ فقط بناء على معالجة آلية لمعطيات يكون الغرض منها تحديد صفات الشخص المعني أو تقييم بعض جوانب شخصيته.

ولا تعتبر قرارات متخذة فقط بناء على معالجة آلية القرارات المتخذة في إطار إبرام عقد أو تنفيذه والتي يكون الشخص المعني قد أتيحت له فيها إمكانية تقديم ملاحظاته وكذا القرارات التي تستجيب لطلبات الشخص المعني".

وإمعاناً في حماية الأفراد في مواجهة الآثار القانونية المترتبة على القرارات الخوارزمية، منح المشرع المغربي بنص المادة ٧/ج من القانون، سالف الذكر، لأصحاب البيانات الحق في معرفة الآلية التي تحكم كل معالجة مؤتممة للبيانات ذات الطابع الذاتي المتعلق به. إذ جاء في هذه المادة على أنه: " يحق للشخص المعني بعد الإدلاء بما يثبت هويته أن يحصل من المسؤول عن المعالجة في فترات معقولة وعلى الفور .

ج- معرفة الأسباب المنطقية التي تحكم كل معالجة آلية للمعطيات ذات الطابع الشخصي المتعلقة به".

---

(1) Margot Kaminski, 'The Right to Explanation, Explained', 2018. See also See Emre Bayamlioglu, 'Contesting Automated Decisions': European Data Protection Law Review 4, no. 4 (2018): 43346, <https://doi.org/10.21552/edpl/2018/4/6>.

وكنا نتمنى على المشرع عند إصدار قوانين حماية البيانات الشخصية العربية عامة، وخاصة التي صدرت مؤخرا "المصري والقطري، والتونسي..." النص على معالجة "خوارزميات الذكاء الاصطناعي" مثلما فعل المشرع المغربي "عبر النص الذي يمكن من تطوير تقنية الذكاء الاصطناعي لمعالجة البيانات القانونية من أجل تمكين تطوير العدالة الخوارزمية في سياق يحترم المبادئ القانونية لحماية البيانات الذاتية لأفراد المجتمع.

## الخاتمة

لقد أدى التطور المتلاحق لأنظمة الذكاء الاصطناعي، من دون تنظيمها وضبطها ضمن الأطر الأخلاقية والتشريعية إلى انتهاكات مروعة في مجال حقوق الإنسان، حيث أدى التحيز الخوارزمي إلى حرمان فئات من حقها في العمل على أساس جنسهم وعرقهم ودينهم وتاريخهم... إلخ، وهو ما يتنافى مع ما دعت إليه الشريعة الإسلامية والقوانين الوضعية، وعلى إثر هذا خلصت الدراسة إلى جملة من النتائج والتوصيات هي في مجملها كالآتي:

### أولاً: النتائج

- التحيز الخوارزمي يعني وجود انحياز في خوارزميات أنظمة الذكاء الاصطناعي، فيتم الحصول على نتائج تتسم بالتحيز لحساب طائفة على أخرى، وهذا بسبب البيانات المتحيزة التي تم اعتمادها أثناء مرحلة التعلم الآلي.
- التحيز الخوارزمي هو امتداد لتحيز وعنصرية البشر المبرمجين.
- تتعدد الأسباب التي تدفع بالخوارزميات إلى التحيز منها ما هو متعلق بتحيز البيانات، والبعض الآخر متعلق بوجود عيب فيها، وغيرها متعلق بفكرة التعميمات.
- يتجسد التحيز الخوارزمي في أنظمة الذكاء الاصطناعي على ثلاثة أنواع هي التحيز التفاعلي، والتحيز الكامن والتحيز الاختياري.
- أدى التحيز الخوارزمي في أنظمة الذكاء الاصطناعي إلى انتهاكات مروعة في مجال حقوق الإنسان، لاسيما الحق في العمل، حيث حرم فئات من فرصة الحصول على العمل بناء على عرقهم وجنسهم ودينهم.
- نظرا للخروقات والانتهاكات التي أحدثتها أنظمة الذكاء الاصطناعي المنحازة في قراراتها في مجال حقوق الإنسان، قام مجلس الشيوخ الأمريكي سنة 2019م بوضع مشروع قانون المساءلة القانونية للخوارزمية.
- قيام منظمة اليونسكو سنة 2021م بسن توصية خاصة متعلقة بأخلاقيات الذكاء الاصطناعي.
- من الآليات المقترحة للحد من تحيز أنظمة الذكاء الاصطناعي على وجه العموم اختيار مجموعة البيانات، وتنويع الفريق، وتقليل الإقصاء... إلخ، ومن الآليات الخاصة للحد من تحيزها في مجال العمل وضع الإجراءات القانونية والتنظيمية، وتدريب أنظمة التوظيف الآلي، ودمج الإشراف البشري في عملية التوظيف ومراقبة وتقييم أنظمة التوظيف الآلي.

## ثانياً: التوصيات:

- يمكن الحد من التحيز الخوارزمي تهدف من خلال تعزيز التصميم العادل واستخدام الأنظمة الآلية المبادئ التي حددها OSTP<sup>(1)</sup> والتي تنص على يجب على مصممي ومطوري مستخدمي الأنظمة الآلية اتخاذ تدابير استباقية ومستمرة لحماية الأفراد والمجتمعات من التمييز الخوارزمي واستخدام الأنظمة وتصميمها بطريقة عادلة.
- يجب أن تشمل هذه الحماية تقييمات المساواة الاستباقية كجزء من تصميم النظام، واستخدام البيانات التمثيلية والحماية من الوكلاء للميزات الديموغرافية، وضمان إمكانية الوصول للأشخاص ذوي الإعاقة في التصميم والتطوير، واختبار التفاوت قبل النشر والمستمر والتخفيف منه، والإشراف التنظيمي الواضح.
- ويجب إجراء تقييم مستقل وإعداد تقارير بلغة واضحة في شكل تقييم أثر خوارزمي، بما في ذلك نتائج اختبار التفاوت ومعلومات التخفيف، ويجب نشرها كلما أمكن ذلك لتأكيد هذه الحماية.
- يعين على أرباب العمل والشركات المعتمدة على أنظمة التوظيف الآلي في دراسة ملفات الأفراد المتقدمين للحصول على العمل محاولة اكتشاف الخوارزميات المنحازة والقيام بتحديثها، إلى جانب إشراك العنصر البشري في عملية انتقاء الملفات.
- يجب على صناع القرار تقييم التكنولوجيا بناءً على معايير موضوعية مثل معدلات الخطأ وأوقات إنجاز المهام ومخاوف الخصوصية، بدلاً من الاعتماد فقط على التصورات الذاتية مثل ردود الفعل العاطفية أو رضا المستخدم.
- ينبغي أن تساعد الاستراتيجيات الإضافية، مثل تكوين الفريق المتنوع، وبحوث المستخدم الشاملة، ومصادر البيانات المتنوعة، واكتشاف التحيز والتخفيف، والتدريب على التحيز، والاختبار التكراري والتطوير، ومراجعات الأخلاقيات، الوكالات على تقليل التحيز في التقنيات الجديدة. تشجع هذه الاستراتيجيات التقنيات الأكثر شمولاً وعدالة وفعالية لمجموعات المستخدمين المتنوعة. ملخصات هذه الاستراتيجيات
- ربط الخبراء في مجال حقوق الإنسان مع الخبراء في مجال تكنولوجيات الذكاء الاصطناعي.
- تصحيح "الخوارزميات الموجودة بشكل يتلاءم مع التطور البشري والمساواة في شتى الحقوق.
- بذل الجهود الدولية والداخلية من أجل تعزيز أخلاقيات الذكاء الاصطناعي.
- يتوجب على الدول التي تتبنى فكرة الذكاء الاصطناعي أن تقوم بتضمين قوانينها الداخلية جملة من النصوص الزجرية لكل من يعتمد إنتاج أنظمة تتصف بالعنصرية.

<sup>1</sup> – Office of Science and Technology Policy <https://www.whitehouse.gov/ostp/>

- دمج مراجعة الأخلاقيات في دورة حياة تطوير التكنولوجيا للكشف عن المعضلات الأخلاقية المحتملة أو التحيزات التي ربما تم تجاهلها. يجب أن تتضمن مراجعة الأخلاقيات خبراء من تخصصات متنوعة. ضمان تضمين المبادئ التوجيهية والمبادئ الأخلاقية، التي تشمل الإنصاف والمساءلة والشفافية والخصوصية، في عملية التطوير منذ البداية.
- يجب أيضا على التدريب المختصين الحد من التحيز من خلال تقديم التعليمات والتوجيه لفرق الأعمال والتطوير فيما يتعلق بأهمية التعرف على التحيزات ومعالجتها، إلى جانب استراتيجيات فعالة للتخفيف. يساعد هذا التدريب على تنمية ثقافة الوعي بالتحيز.

## قائمة المصادر والمراجع

### أولاً: المصادر

- التوصية الخاصة بأخلاقيات الذكاء الاصطناعي منشورة على موقع منظمة اليونسكو  
[/https://unesdoc.unesco.org](https://unesdoc.unesco.org)
- معجم البيانات والذكاء الاصطناعي انجليزي-عربي، ط1، الهيئة السعودية للبيانات والذكاء الاصطناعي، مجمع الملك سلمان العالمي للغة العربية، 2022م ثانياً: المراجع باللغة العربية  
١- المقالات الأكاديمية
- هايدي عيسى حسن علي حسن: حقوق الإنسان في عصر الذكاء الاصطناعي " معطيات ورؤى وحلول"، مجلة الشريعة والقانون، ع 85، يناير 2021م.
- ٢- مقالات الإنترنت
- دان رايلي: للتغلب على التحيز في الذكاء الاصطناعي يجب ألا الاعتراف بوجوده، منشور على الموقع  
[/https://fortunearabia.com](https://fortunearabia.com)
- مصطفى السيد حسين: عصرنة الذكاء الاصطناعي... لماذا يبدو أصدقاءك بملامح أوروبية عند استخدامهم تطبيقات وفلاتر الذكاء الاصطناعي؟ مقال منشور على الموقع  
<https://arabicpost.net>
- نور علوان، حفصة جودة: تحيز العالم الرقمي... عندما تشجعنا محركات البحث على العنصرية مقال منشور  
<https://www.noonpost.com>
- سكيئة السمرة: الذكاء الاصطناعي بين جدلية التفوق والتحيز... هل السبب في الخوارزميات، مقال منشور على الموقع  
<https://www.annahar.com/arabic/section/4>
- سوزانا شلينبرج: التحيز الخوارزمي وتصعيد اللامساواة، ترجمة: زياد الحازمي، تحرير: ت ركي طوهري، مقال منشور على  
<https://atharah.net/how-biased-algorithms-perpetuate/inequality>
- طه الراوي: عنصرية التكنولوجيا... عن أسباب وأشكال التحيز الرقمي، مقال منشور على الموقع  
. <https://www.noonpost.com/content/37283>
- محمد فتحي: الذكاء الاصطناعي ينافس الإنسان في الانحياز مقال منشور على  
[/https://www.alkhaleej.ae](https://www.alkhaleej.ae)
- كيف يمكن التغلب على تحيز الذكاء الاصطناعي... تقنيات وأدوات، مقال منشور على

<https://masaar.net/ar>

- تصميم الذكاء الاصطناعي بحيث يكون منصفاً مقال منشور على الموقع

<https://arabicedition.nature.com>

- الذكاء الاصطناعي متهم بالعنصرية، مقال منشور على الموقع

[/https://www.alaraby.co](https://www.alaraby.co).

- الذكاء الاصطناعي: تحقيق التوازن بين الابتكار وحقوق الانسان مقال منشور على

<https://masaar.net/ar> م. 22/05/2023 بتاريخ

- كيف يمكن لتطبيقات الذكاء الاصطناعي أن تكون متحيزة مقال منشور على الموقع

<https://masaar.net/ar>

### ثالثاً: المراجع الأجنبية

- Lindsey Andersen Human Rights in the Age of Artificial: Intelligence  
<https://www.accessnow.org/wp-content/uploads/2018/11/AI-and-Human-Rights.pdf>
- Booker, Wyden, Clarke Introduce Bill Requiring Companies To Target Bias In Corporate Algorithms, published on  
<https://www.booker.senate.gov/>, 10/04/2019.