مجلة العلوم الاجتماعية والإنسانية
Journal of Social Sciences and Humanities
المجلد: 15/ العـدد: 02(2025)،
ص 718 - 736

النزعة العو اقبية في أخلاقيات الذكاء الاصطناعي

Consequentialism in the Ethics of Artificial Intelligence

د. بوزينة حنان[*1]

HANANE BOUZINA

[1-]أستاذ محاضر ب جامعة الشهيد حمه لخضر جامعة الوادي

hanane-bouzina@univ-eloued.dz

تاريخ الاستلام: 2025/07/09 تاريخ القبول: 2025/11/24 تاريخ النشر: 2025/12/15

ملخص:

يعتبر هذا البحث محاولة لتوضيح النهج الأخلاقي العواقبي الذي تقوم عليه أخلاقيات الذكاء الاصطناعي، الأمر الذي أفرز فريقين مختلفين حول نجاعة هذا النهج الأخلاقي في تسيير النظام الخوارزمي وتحقيق الرفاهية المرجوة من عدمه، مما جعلنا نقف عند الأسباب التي جعلت كل فريق يتحفظ برأيه. وقبل كل ذلك عرضنا لمحة دلالية على الأخلاق المعيارية التي تعتبر الأخلاق العواقبية قسم من أقسامها إلى جانب الأخلاق الواجبية وأخلاق الفضيلة، حتى نسهل على القارئ فهم اللاحق من المقال. وقد قادتنا طبيعة الموضوع إلى اعتماد المنهج التحليلي الوصفي والمنهج النقدي من أجل دراسة عميقة وبسيرة لأخلاقيات الذكاء الاصطناعي التي أصبحت مشكلة العصر بعد سيادة التكنولوجيا. وقد توصلنا في نهاية البحث إلى جملة من النتائج كانت أهمها تلك المتمثلة في عدم كفاية النهج العواقبي الأخلاقي في تسيير أنظمة الذكاء الاصطناعي .

الكلمات المفتاحية: الأخلاق، الأخلاق المعيارية، الأخلاق العواقبية، أخلاقيات الذكاء الاصطناعي.

Abstract:

This research attempts to clarify the consequentialist ethical approach underlying AI ethics. This approach has produced two different groups regarding the effectiveness of this

[*] المؤلف المرسل: بوزينة حنان ، الإيميل: hanane-bouzina@univ-eloued.dz

CC BY NC

ethical approach in managing algorithmic systems and achieving the desired well-being. This has led us to examine the reasons behind each group's reservations. Before all that, we provided a semantic overview of normative ethics, of which consequentialism is a subset, along with deontological ethics and virtue ethics, to facilitate the reader's understanding of the rest of the article. The nature of the topic led us to adopt a descriptive analytical approach and a critical approach for a deep and accessible study of the ethics of artificial intelligence, which has become a contemporary problem following the rise of technology. At the end of the research, we reached a number of conclusions, the most important of which was the inadequacy of the ethical consequentialist approach to managing artificial intelligence systems.

Keywords : Ethics, Normative ethics, Consequential ethics, Ethics of artificial intelligence.

Résumé:

Cette recherche tente de clarifier l'approche éthique conséquentialiste qui sous-tend l'éthique de l'IA. Cette approche a donné lieu à deux groupes de points de vue différents quant à son efficacité dans la gestion des systèmes algorithmiques et l'atteinte du bien-être souhaité. Cela nous a conduits à examiner les raisons des réserves de chaque groupe. Avant tout cela, nous avons présenté un aperçu sémantique de l'éthique normative, dont le conséquentialisme est un sous-ensemble, ainsi que de l'éthique déontologique et de l'éthique de la vertu, afin de faciliter la compréhension du reste de l'article. La nature du sujet nous a conduits à adopter une approche analytique descriptive et une approche critique pour une étude approfondie et accessible de l'éthique de l'intelligence artificielle, devenue un problème contemporain suite à l'essor des technologies. À l'issue de la recherche, nous sommes parvenus à un certain nombre de conclusions, dont la plus importante était l'inadéquation de l'approche conséquentialiste éthique dans la gestion des systèmes d'intelligence artificielle.

Les mots clés : Éthique, Éthique normative, Consequentialist ethics , Éthique de l'intelligence artificielle.

مقدمة

تُشكّل الأخلاق ركنا أساسيًا في الحياة الاجتماعية حيث تؤدي الدور المنظم والموجه للأفعال البشرية، ويصدّع غيابها المجتمع وقد يؤدي به إلى الانهيار. فهي توجه الإنسان إلى التصرف الحسن بنسب إرسالية وفق ما تقتضيه غاياته السامية، ونسب إمساكية تمنعه عن كل ما ينأى به ويخالف ما تقتضيه حقيقته. وهي بهذا تمثّل مجموعة الزامات تحدد ما ينبغي فعله وما لا ينبغي فعله. وقد انقسمت من حيث المبحث إلى ثلاثة أقسام متباينة: الأخلاق المعيارية التي توفر القواعد الأخلاقية العامة الضابطة لسلوكنا بخصوص ما ينبغي القيام به أو تركه، بالإضافة إلى الأخلاق التطبيقية التي تهتم بكيفية التعامل مع الظروف الخاصة بحياتنا العملية على سبيل المثال لا الحصر المسائل الطبية المتعلقة بالموت الرحيم والمسائل الاجتماعية كتحريم أكل بعض اللحوم، وآخر قسم هو الميتا-أخلاق التي تقع في الرتبة الثانية من البحوث الأخلاقية التحقيقية لأنها لا تبحث في مجلى الفعل الحسن أو القبيح، وإنما الذي يهمها مفهوم الحسن والخير، فتتناوله من المحكي الدلالي من أفق الاستعمال والفهم، والمحكي الأنطولوجي حقيقة تمثّل الخصائص الأخلاقية ذهنيًا وواقعيًا، والمحكي الابستيمي الذي تعالج فيه بيان وجود معرفة أخلاقية من عدمها.

وقد جرت العادة أن تتعامل الأخلاق مع المواقف الأخلاقية التي لها علاقة كيانيّة بالواقع، كأن نقول "الكذب قبيح"، فالحكم المتضمن في متن هذه القضية جرى تطبيقه على الممارسات التي تتم في صقع الواقع. ومع تسارع وتيرة الاختراعات والمكتشفات التي صاحبت ميلاد العصر الحديث، والتي طالت الجانب التكنولوجي، أين استطاع الإنسان بفكره أن يخلق عالمًا آخر يغاير العالم الخارجي من ناحية البعد، فإذا تحدثنا عن البعد البشري فنحن في متن الواقع المادي، أما إذا تحدثنا من منطلق البعد التكنولوجي فنحن في متن الواقع الافتراضي، الذي اصطلح على تسميته " العالم السيبراني" الذي هو عبارة عن شبكة عالمية رقمية، استطاعت أن تُغيّرَ من طريقة عيشنا واختزال سبلها سواء أكان ذلك من الجهة التواصل (اختصار المسافة)، أم من جهة الاستيعاب (تسهيل الفهم) وغيرها من التسهيلات التي طرحت العديد من الإشكاليات الأخلاقية منها التي تتعلق بالموثوقية والخصوصية والأمن. فهل ستتمكن الأخلاق بأقسامها الثلاثة من جعل هذا الفضاء الرقمي أكثر أمانًا مثلما هو الحال مع أخلاقيات المجتمع البشري. بعد أن عرفت البيئة الافتراضية نموًا متسارعًا على مستوى الممارسات والتقنيات المرتبطة بالذكاء الاصطناعي، الذي تعددت استخداماته لتشمل العديد من القطاعات الصحية والتعليمية والتربوية والدفاعية وحتى الترفيهية... من طريق إنشاء وتشغيل أنظمة ذات نماذج وخوارزميات بأساليب متطورة تساعد على تحسين جودة المعلومات.

ونحن بحديثنا هذا، لسنا نسعى إلى تعزيز قدرات الذكاء الاصطناعي، بقدر ما نطمح إلى تبيان مدى التزام

تلك التقنيات الرقمية للقواعد الأخلاقية. وقد اخترنا أن ندرسها من باب الأخلاقيات المعيارية وبالضبط الأخلاق العواقبية التي تحدد قيمة الفعل الأخلاقي من خلال نتيجته، إذ لا يهمها طبيعة الفعل إن كان حسنًا أم قبيحًا. وعليه نطرح السؤال التالي: هل تم مراعاة القواعد الأخلاقية في أنظمة الذكاء الاصطناعي أم أنه تم فقط التركيز على عواقبيته التي توفر الخير الذاتي؟ أو بتعبير آخر: هل يمكن أن تكون أخلاقيات الذكاء الاصطناعي ذات نزعة عواقبية؟

أولا: لمحة معرفية حول الأخلاقيات العو اقبية

تقع الأخلاق النفعية في قسم الأخلاق المعيارية إلى جانب أخلاق الواجب Deontological ethics التي تحدد قيمة الفعل بمدى التزام الإنسان بالواجب الأخلاقي Moral duty (لا تسرق، لا تكذب، لا تقتل...) بصرف النظر عن الموقف الذي هم في متنه، لأنّ القانون الأخلاقي يمنع ذلك، فلن ينفعه حبه لشخصٍ ما ولا ولاؤه لوطنه ولا حتى معزّة نفسه في تجاوز الواجب الأخلاقي؛ لأنه مطالب فقط بأن يحترم شخصي القواعد العملية مثل الوفاء بالوعود وعدم الكذب وعدم إذلال الآخرين ... مهما كانت النتائج من وجهة نظر محايدة أو غير شخصية (مونيك كانتو سبيربير، روفين أدجيان، 2008، صفحة 79) فالواجبية الكانطية لا تعترف بالتجاوزات العاطفية التي تحكمها ميولات الحب والنخوة والقرابة...الخ. لنفترض مثلا أنّك محاصر عند العدو وهددوك بالايشاء بمكان الجنود، فأنت كفاعل أخلاقي عادي لن تشي بذلك ولو كان ذلك على حساب حياتك، ولكن هذا غير جائز من الرؤية الواجبية الكانطية لأنّك خرقت الواجب الذي يأمرك بعدم الكذب مهما كان الظرف (لا تكذب) وبالتالي ستصدق وتشي بمكانهم، لأن مخالفة الواجب الأخلاقي تعتبر من عظائم الكبائر.

ويعود حرص كانط على الالتزام الكلي بالواجب الأخلاقي، إلى طموحه في التأسيس لأخلاقيات عالمية، وتلك التجاوزات المشروطة بالميول النفسانية تجعل من الواجب الأخلاقي ذاتانيًا. وتمسي الأخلاق بمقتضى ذلك محكي نفسانيًا والقانون الأخلاقي تكليفًا ذاتيًا. وعليه فالقيمة الأخلاقية في متن الواجبية يحددها الفعل الحسن الذي حدده الواجب ورشحته الإرادة الحسنة good will التي اعتبرها كانط مِلاك الأفعال الحسنة، تجلى ذلك بوضوح في قوله: الحقيقة الوحيدة التي يجب الإيمان بها تتمثّل بالخضوع بحتمية إلى الواجب الأخلاقي المتعين بالإرادة...التي تحرك الفاعل الأخلاقي قبل قيامه بفعل ما، إلى أن يفكر في حدودها، ويقوم بما يستطيع أن يقدر عليه (Emmanuel Kant, 1888, page 30)، بمعنى أنّ الفاعل الأخلاقي يتحرك بموجب ما تشرعه إرادته الخيّرة المفطورة في هندسته التكوينية، والجاذبة إلى كل ما هو خير، بحكم النية الصافية المتعينة في خلقته. وقد تتساءل لماذا لم يقع اختيار كانط على الدين بحكم المعيارية المنزهة؟ السبب لا يكمن في كون كانط يرفض الدين –مثلما فعل معظم فلاسفة التنوير-بل في كونه يرغب في التأسيس لأخلاق غير مشروطة لا بالخير الذاتاني (حظوظ النفس من حب وصداقة وقرابة...) ولا الخير

الموضوعي (مرضاة الإله من ثواب وفوز بالجنة). وهذا ما جعله يلجأ إلى الإرادة الحسنة التي تعمل وفقَ الواجب وغير المشروطة بشيء، وهذا ما جعلها بهية ونقية وبالتالي فهي تعتبر بشكل عام السمة الأساسية للحياة الأخلاقية لصفائها وبهائها وبالتالي يجب عدم الخلط بينها وبين الميل الذاتي حتى لو كان ميلاً للشفقة والإحسان ...فهي لا تستمد صلاحيتها من عواقب الأفعال، كما أن صوابية الفعل تكمن فقط في الإرادة الخيّرة (Josiane Boulad-Ayoub, pages 84-85) لنفترض أنّك في موقف كنت فيه شاهد على حادثة، فإنّك ستصدق لأنّ إرادتك الحسنة دفعتك إلى قول الصدق، فحينها أنت لم تفكر في الحسنات (مرجعية دينية) ولا في المصالح (مدح الناس)، وإنّما صدقت لأنّ ذلك هو الفعل الصائب المترشح من نيتك الصافية، وعدا ذلك يبقى تفكيرًا نتائجيًا حركته حظوظ النّفس ولا شأن للإرادة الحسنة في ذلك.

نأتي الآن إلى بيان القسيم الثاني من قسم الأخلاق المعيارية وتمثله أخلاق الفضيلة Virtue ethics، حيث تتحدد القيمة الأخلاقية ضمن هذه الرؤية بالفاعل الفاضل بمقتضى الموقف الذي هو فيه، حيث يتصرّف على مقتضاه –الموقف- بالطريقة التي يراها صائبة فهو الوحيد الذي يمتلك الحكمة العملية والوحيد القادر على تقييم ذلك الموقف الذي يعيشه، وبالتالي إنّه غير ملزم باحترام قاعدة أخلاقية (يجب أو لا يجب) خارجة عن رهانات الوضع الملموس، لأنه يدرك تمامًا أنّه أمام العديد من الاختيارات وسيختار الفعل الذي يورثه ويورث لغيره السعادة (Konstantin Büchler, 2014, pages 5-6) لذلك نجد أخلاقيات الفضيلة تعيب على الرؤى الأخلاقية التي تجعل من الفعل الأخلاقي الحسن شرطًا للقيمة الحسنة، فهي بهذا الشرط تُغَيِّبُ ذات الفاعل بوصفه العنصر الفعال في الموقف الأخلاقي، إذ حصرته بمجموعة معايير التي لا تصلح في جميع المواقف. فطبيعة الإنسان تميل إلى التعلم من خلال الملاحظة والتجربة، مما يجعله يكتسب سلوكات فاضلة، وليس التعلم فقط من خلال إخضاعه لمجموعة قواعد يمتثل لها. كما أنّه بدلاً من التركيز على العواقب الجيدة يجب التفكير في طريقة تجعل من الفاعلين فضلاء وصالحين (David Massey, pages 1-2) .

وتمثل الأخلاق العواقبية Consequential ethics القسيم الثالث من قسم الأخلاق المعيارية –وقد تعمدنا تأخيره لكونه يمسّ موضوع بحثنا-. إنّ القيمة الأخلاقية في متن هذا المحكي تحدده نتيجة الفعل بغض النظر عن جوهر الفعل أكان حسنًا أم قبيحًا. فالنتيجة هي الملاك الأول والأخير في الموقف الأخلاقي ففي بعض المواقف يجد الفاعل الأخلاقي نفسه مضطرًا للكذب إذا كان يعتقد أنّ ذلك ضروري لتحقيق أفضل النتائج (Section B, page118) ففي حالة الحرب –على سبيل المثال لا الحصر- قد يضطر المرء إلى الكذب لصالح وطنه وقد يضطر إلى القتل للدفاع عن نفسه وما أشبه ذلك من الاستثناءات التي تجاوزتها أخلاقيات الواجب التي تلزمك بالالتزام بالواجب مهما قست الظروف.

وتعتبر المدرسة النفعانية من المدارس الأخلاقية التي ترى أنّ قيمة الفعل تحددّها نتيجته التي تورثنا

الشعور بأكبر قدر ممكن من السعادة (Section B, page 119) . حيث يرى كل من الفيلسوف بنثام جريمي والفيلسوف جان ستيوارت مل – من روّاد المدرسة - أنّ المِلاك الأساسي للأفعال البشرية يتمثّل بفائدتها ومقدار ما تضفيه من السعادة على البشر وتخفف عنهم البؤس والمعاناة (Le bonheur, pages 9-10) . مما يعني أنّ العواقبية جعلت من السعادة الغاية القصوى للفاعل الأخلاقي مهما كانت الوسيلة التي استعان بها الفعل. وهذا يذكرنا بالفكرة الميكيافللية التي مفادها أنّ الغاية تبرر الوسيلة.

تأسيسًا على ذلك، يرى بنثام أنّه لتحقيق السعادة يجب تغيير المصدر الذي يشرع القواعد ويجعل منها الزامات، وهو في حديثه هذا يقصد الدين الذي كان يمثّل المرجع الأساسي في الأخلاق التقليدية، وهو يُنَحيه جانبا لأنّه –حسب منظوره- يقوم على مبدأ الزهد والخوف المبالغ فيه من العقاب الأخروي النابع عن الهوى الخرافي، الخوف من غضب الإله المنتقم...وقد ألقت هذه الخرافات الدينية بسوادها الأعظم على حياة الإنسان، فجعلته مثالاً للبؤس والشقاء والمذلة نتيجة الخوف الذي لا مبرر له... وبالتالي إنّ الواجب الأخلاقي الذي مرجعه الدين، جعل البؤس أمرًا مرغوبًا فيه (J.Bentham, pages 15-20-21). ومادام الدين يحرّك الفاعل الأخلاقي بدافع الخوف فإنّه لا يصلح كقاعدة حياتية، لذلك استبدله بالمنفعة التي يراها قاعدة عملية تخدم واقع المرء أفضل من المتعاليات التي قيدته بمسمى الدين الحق وكبحت إرادته الأمر الذي قوَّض حريته في تحصيل سعادته.

تأتيًا للسابق ذكره، يمكن القول أن أخلاقيات المنفعة نظرية في السعادة، السعادة العظمى التي تمثل مجموع المتع التي تم طرح الآلام منها، وقد أصبحت اليوم –السعادة- هدفا عمليًا مقبولاً بشكلٍ عام كما في الماضي التي يمكن تمسّ أكبر عدد من الناس (Jean-Pierre Cléro, 2011, page 465)).

إنّ الذي يهمنا من هذا الطرح البسيط الذي عرضناه بإيجاز تقسيمات الأخلاق المعيارية التي تنتمي إليها الأخلاق العواقبية التي هي محور بحثنا، ولكن في علاقتها بأخلاقيات الذكاء الاصطناعي التي تميز روح عصرنا، وهذا ما يتضمنه المبحث الموالي عرضًا وتفصيلا.

ثانيا: النزعة العو اقبية في أخلاقيات الذكاء الاصطناعي

يمكننا إبراز النزعة العواقبية التي تتضمنها أخلاقيات الذكاء الاصطناعي من خلال مجموع المعايير التي تساهم في سيرورة الذكاء الاصطناعي، وهي كالتالي:

1-المناسبة

يتفق معظم المفكرون على أنّ الأخلاق العواقبية هي المناسبة لأخلاقيات الذكاء الاصطناعي بالنظر إلى باقي التقسيمات الأخلاقية، والسبب في ذلك راجع أولاً إلى طبيعة الذكاء الاصطناعي المستقبلية التي تتكهن

الأحداث البعيدة التي تعمل على جعلها على قدر أعلى من الأفضلية. وثانيًا من جهة الغائية حيث تجهز كل آلة على أداء وظيفة معينة بغاية محددة مسبقًا، وهذا ما أكدّه الذكاء الاصطناعي المعاصر الضيّق الذي يسير على هدي تلك السمتين العمليتين اللتين أورثتهما العواقبية النفعانية (Josiah Della Foresta, 2020, page 01) حيث تسعى أخلاقيات المنفعة -كما تحدثنا عن ذلك سالفًا- إلى تحقيق أكبر عدد من النتائج التي تعود بالنفع والرفاهية على البشر، فهي تضع المستقبل نقطة انطلاقها، وهذا ما لا نجده في الأخلاقية الواجبية التي تحدد قيمة المستقبل انطلاقًا من قيمة الفعل، وكذلك في الأخلاق التطبيقية التي تركّز كذلك على قيمة الفعل في الموقف الأخلاقي.

وخارج تلك المزايا المذكورة، هناك من يرى أن الأخلاقيات التي تقوم على الدعائم المتعالية (الدين، النوايا الحسنة ...) لا تصلح أن تكون نهجًا يهتدي به أخلاقيات الذكاء الاصطناعي لأنّها تضيّق من مساحة الفعل في تحقيق النتائج المرجوة. وهذا ما يذكرنا بوجهة نظر جون راولز حول النظم الأخلاقية المتعالية التي تصلح أن تكون مرشدًا للعلاقات الاجتماعية أكثر من أن تصبح قاعدة عملية لأخلاقيات التقنية. ولهذا تبرز الأخلاق العواقبية القائمة على تحقيق المنفعة كبديل لتلك النظريات الأخلاقية – ليس بالضرورة أن تكون بديلاً متفوقًا- ولكن يمكن أن نعتبرها بديلاً ملائمًا أكثر منهم في تسيير وتحسين تطبيقات الذكاء الاصطناعي، نظرًا لطبيعتها التي تتقارب مع طبيعة التعلم الآلي (Dallas Card, Noah. Smith,2020, pages 1-2)بمعنى أن أخلاقيات الذكاء الاصطناعي ذات طبيعة عملية لذلك تناسبها الوظيفة التي تدعم النتائج الجيدة بغض النظر عن قيمة الفعل إن كان حسنًا أو قبيحًا، ولا تلائمها الوظيفة الارشادية التي تساعد في تقوية الروابط الأخوية بالآخرين لذلك تهتم بقيمة الفعل لأنه محرك التواصل بينهم.

وضمن الطرح نفسه، يتجه فريق آخر من الباحثين إلى القول بأنّ الأخلاق العواقبية هي الوحيدة التي تمتلك المقدرة في إعادة تصميم التكنولوجيات وتصنيعها وتغيير النتائج الغير متوقعة مسبقًا، لأنّها تعمل مسبقًا على ترصد النتائج الجيدة والسيئة – كاحتمال- مما يسمح لها في وضع خطة احتياطية لتفادي ذلك. وهذا ما يضمن تطوير الذكاء الاصطناعي بدلاً من حدوث الأسوأ(Stefan Buijsman and others, 2023, page 11) ولعل ما نشهده من تطورات متعاقبة لتطور الآلة وتفوقها في الأعمال التي كان يقوم بها الإنسان، ومؤخرًا تم إشراكها في عالم البيولوجيا، يثبت سالف الحديث، وهذا راجع إلى عدم تقيد أخلاقيات العواقب بالوازع الديني مما يجعلها تتوقع مسبقًا الحالات الجيدة والحالات السيئة، فعلى سبيل المثال لا الحصر، نأخذ فعل "القتل" مثالاً لتوضيح ذلك، ونفترض برمجة آلة لتخوض الحرب، فإن أردنا تحقيق نتائج عظيمة فإننا لن نكترث إن قتلت طفلاً أو حاملاً أو أعزلاً، بينما إذا أردنا أن تقوم بوظيفتها من هدي الأخلاق الإلهية أو حتى الواجبية، فإن برمجتها ستعمل على العدل عن القيام بما سلف ذكره في المقترح الأول. وهذا ما يذكرنا بالفيلم الأجنبي The Terminatorالذي يصور العالم بعد سيادة الآلة، في البدء كانت تُقْدِمُ الآلة المبرمجة مسبقًا على قتل أي عابر ولكن بعد أنسنة فعلها تغير الوضع لتقتل فقط

من يحاول الاعتداء.

ولما كان حفظ الأمن وضمان سلامة المواطنين من واجب المؤسسات، فإن اللجوء إلى تطبيق الأخلاق العواقبية في الذكاء الاصطناعي، لأنها -العواقبية- تعمل على تقييم العواقب طويلة المدى لاستخدام البيانات الشخصية على المجتمع والبيئة والأفراد. اعتماد على الممارسات التي تحقق أقصى قدر من الفوائد مع تقليل الأضرار المحتملة (Ayoub El Azami El Idrissi et autres, 2024, page 31).

2-النتائجية

إنّ الذي يحدد القيمة الأخلاقية في الأخلاق العواقبية هو النتيجة سواء أكانت مترتبة على فعل حسن أم قبيح، لأن الذي يهم العواقبيون النتيجة الحسنة التي توفر أكبر قدر من الرفاهية، وهذا ما يسمى بمبدأ العواقبية في التصرف، حيث يركزون فقط على التصرف الذي يُوَّفر نتائج أفضل، حيث يمكننا حساب القيمة الأخلاقية في متن هذه الرؤية كالتالي: ق= (النتيجة -الفعل) بمعنى تأخير قيمة الفعل على العواقب بحيث تبدو العواقب أشد أهمية من الفعل لأن المهم –هنا- حساب فقط جميع العواقب سواء البعيدة منها أو القريبة، مهما كان التصرف المتخذ إزاء ذلك، وهذا يساهم – حسب منظورهم - في الكيفية التي يتم بها تطوير النتائج في لحظة زمنية معينة (Dallas Card, Noah. Smith, 2020, page 03) مما يعني أن العواقبية تهتم بثلاث نقاط رئيسية: النتيجة، التطور، الزمن، فمن خلال هذه المعطيات يمكن توفير الرفاهية وتحقيق النجاح.

وبالنظر إلى أخلاقيات الذكاء الاصطناعي، نرى أنها تأسست على مثل هذا الحساب الذي يهتم فقط بالنتائج المحصلة، حيث نرى أنّ عالم الكمبيوتر يتصرف بشكل أخلاقي إذا قام ببرمجة أو تدريب الآلة بحيث تقتل أقل عدد من الناس؛ وأن المؤسسات (على سبيل الدولة) الجيدة هي التي لديها مخاوف أخلاقية، فتلجأ دومًا إلى انتهاج العواقبية في تطوير المركبات الذاتية القيادة حتى تتنبأ مسبقًا بالنتائج التي يمكن أن تمسّ المواطنين، فتعمل بذلك على وضع كل الاحتياطات بناء على الحسابات المسبقة التي تحافظ على المواطنين وتقلل من أخطار حوادث السير(Allison Marchildon et autres, 2023, page 11).

تتجلّى النزعة النتائجية في معضلة "الترام" وهي تجربة تجريدية توضح المعضلات التي يمكن أن تواجه السيارة ذاتية القيادة أثناء سيرها –خاصة وأن الفاعل الأخلاقي –هنا- ليس شخصًا بل آلة- مما يستلزم برمجتها على أساس تحقيق أفضل النتائج في مثل تلك المواقف التي قد تقتضيها الصدفة. على سبيل المثال، في الحالة التي يمكن فيها وقوع الحادث أمر لا مفر منه، وحيث يمكن للمركبة أن تستمر في سيرها وتقتل (س) من الأشخاص ولنفترض أن عددهم خمسة أو تنحرف إلى وجهة أخرى وتقتل (ع) من

الأشخاص ولنفترض أن عددهم اثنان، فإنها ستنحرف لتحقق النتيجة الأفضل ولن تقتل أكبر عدد. وإنّ في انحرافها معضلة أخلاقية أخرى لأنه في كلتا الحالتين فعل للقتل (Allison Marchildon et autres, 2023, page, 10)غير أننا لسنا هنا بصدد التفصل في تلك المعضلة لأننا سنلحق بيان ذلك في لاحق المقال.

3-المنفعة

أوضحنا فيما سبق، أنّ الأخلاق العواقبية تقوم أساسًا على نظرية المنفعة التي تحدد قيمة الفعل الأخلاقي بما حَصَّله من منفعة، وقد لاقت هذه الضابطة الأخلاقية رواجًا عند القائمين على تسيير وتطوير الذكاء الاصطناعي، حيث أظهر أندرسون وآخرون كيف يمكن تنفيذ النفعية القائمة على فعل اللذة لحساب الفعل الصحيح بناءً على أفضل العواقب بطريقة متسقة وكاملة وعملية. فحسبهم أن حساب المنفعة هو الأكثر ملائمة لذكاء الآلة من الحساب الأخلاقي البشري، وكذلك أكثر موثوقية (Josiah Della Foresta, 2020, pages 3-4) وبذكر مصطلح اللّذة الأخلاقية نستحضر الأخلاقية الأبيقورية، التي مهدت الطريق للعواقبية الحديثة حين اعتمدت اللّذة معيار الخير الأخلاقي، فكل ما يورث النّفس المتعة يعتبر حسنًا والعكس (ثنائية اللذة والألم). وخير الآلة يقع أولا في توفيرها الجهد على الإنسان، وثانيا تحقيق سعادة الفرد والتقويض من دائرة بؤسه ومعاناته لذلك اعتبر كل من جريمي بنثام وجون ستيوارت مل أنّ الملاك الأساسي للأفعال البشرية يتمثّل بفائدتها ومقدار ما تضفيه من السعادة على البشر وتخفف عنهم البؤس والمعاناة(Le bonheur, pages 9-10) ومن أمثلة ذلك نذكر مواقع التواصل الاجتماعي التي بفضلها أصبح العالم قرية صغيرة، حيث وفرت تعب السفر واختزلته في آلات ذكية تعمل وفق نظام خوارزمي، يحفظ الروابط الاجتماعية والعملية...إلخ

وضمن المحكي نفسه، هناك من يرى أن النفعية كنهج يصلح اعتماده كمبدأ أخلاقي للذكاء الاصطناعي، كونها تحفظ السلامة عند حصول أي ظرف طارئ، من دون خلق صراع أخلاقي، فمثلاً بالنسبة للسيارة ذاتية القيادة –وهو مثال سبق استخدامه- لو افترضنا أن السيارة مبرمجة على ترك الأمور تسير كما كُتب لها أن تكون، فإن الضرر سيكون أكثر خاصة إن كان مكان الحادث طريق مزدحم. لذلك فالنفعانية مفيدة لذكاء الآلة من باب الحفاظ على سلامة أكبر عدد من المواطنين (Josiah Della Foresta, 2020, pages 8-9) بمعنى أن النفعية تُسبق – أحيانا - الصالح العام على الصالح الخاص، وهذا ما يذكرنا بنظرية الصواب الموضوعي عند جون ستيوارت مل، والتي تحكي أنه يجوز التضحية بالخاص إذا كان ذلك يخدم العام، على سبيل المثال لا الحصر، إنّ الاقدام على قتل شخصٍ يُلحق الضرر بالآخرين، يترتب عنه عواقب طيبة تنفع الجميع. وهذا ما أطلق عليه هو –مل- بنظرية الصواب الموضوعي في مقابل الالتزام الأخلاقي الذي يُقَبِحُ فعل القتل (Michael Schefczyk) .

هناك من يعتبر أن الأخلاق النفعانية هي الرؤية البديلة الناجحة التي تضع مساحات لتوقع الأحداث، حيث يتم اختيار الإجراءات بناءً على حالة العالم، فهي لا تنطلق من المتعالي بل من المحايثة الواقعية، وهذا ما يناسب أخلاقيات الآلة، فقد يفشل الفاعل الأخلاقي البشري في اتخاذ القرار الصحيح إزاء موقف مفاجئ وتفلح في المقابل الآلة وبدقة حسبما ما ذهب إليه بعض الباحثين في مجال الذكاء الاصطناعي (Dallas Card, Noah Smith, 2023, page 03).

وقد تم اختيار هذه النظرية –النفعانية- لتسيير الذكاء الاصطناعي لأنّ هذا الأخير يعمل بكميات كبيرة من بيانات المستخدمين، مما يساعده على تحقيق الصالح العام للمجتمع. وبالتالي فإنّ نظرية النفعية هي واحدة من أكثر الأساليب شيوعًا لاتخاذ القرارات الأخلاقية، وخاصة القرارات ذات العواقب التي تهتم بمجموعات كبيرة من النّاس، حيث تُعلمنا النفعانية مسبقًا أن نترقب نسب الخير والشر المترشحة عن أفعالنا (Alexander Mitov, 2025, page 05).

تأتيا مما سبق ذكره، هناك من يلخص صلاحية الأخلاق النفعانية للذكاء الاصطناعي في ثلاث نقاط (Alexander Mitov, 2025, page 05) أولا: توفر النفعية أساسًا واضحًا ومباشرًا لصياغة واختبار السياسات. وفقًا للمعايير النفعية، حيث تكون السياسة أو القرار أو الإجراء التنظيمي حسنا إذا كانت تعزز الصالح العام أكثر من أي بديل آخر. ثانيا: توفر النفعية وسيلة موضوعية وجاذبة لحل الصراعات المتعلقة بالمصالح الذاتية. حيث تتناقض هذه السمة -من سمات النفعية- بشكل كبير مع الأنانية، التي تبدو غير قادرة على حل مثل هذه الصراعات... وهكذا، يتخذ الأفراد داخل المنظمات قرارات أخلاقية ويقيمون أفعالهم من خلال اللجوء إلى معيار موحد: الصالح العام. ثالثا: توفر النفعية نهجًا مرنًا وموجهًا نحو تحقيق النتائج في اتخاذ القرارات الأخلاقية... وهذا الجانب من النفعية يمكّن المنظمات من اتخاذ قرارات أخلاقية واقعية وقابلة للتنفيذ.

4-المصلحة:

تعتبر المصلحة إحدى المبادئ المهمة التي تركّز عليها الأخلاق العواقبية كونها حيثية تتماهى مع تحقيق العواقب الجيدة، وهي خاصية كذلك يعمل على هديها وكلاء الذكاء الاصطناعي، حيث يأخذون في حسبانهم المصالح التي تعود عليهم بالنفع، مما يدفعهم إلى مراعاة قيم جميع أصحاب المصالح، كل من أولئك الذين يتفاعلون بشكل مباشر مع نظام الذكاء الاصطناعي وأولئك الذين يتأثرون بشكل غير مباشر باستخدامه. الأمر الذي يتطلب المشاركة النشطة لأصحاب المصالح في استخلاص القيم المختلفة التي تهمهم. (Stefan Buijsman and others, 2023, page 13)

وهم في استخلاصهم لمجموع القيم التي تناسب تقنيات الآلة ينطلقون من الأسفل إلى الأعلى في مقابل

الأطر المعيارية المتعالية التي تأتي من الأعلى إلى الأسفل، حيث يمكن أن تكون –مثلا- خاصية الاستقلالية ذات قيمة في ادراكات العامة من النّاس في حين لا تحمل أي قيمة عند المساهمين الفعالين في مجال التقنية وهم أصحاب المصلحة، لأنهم يعملون ضمن نطاق مقيد بالتكنولوجيا الخوارزمية وليس في مجالٍ واسع، فما يليق بالمعياريات الأخلاقية في المواقف الحياتية المختلفة قد لا يليق بأخلاقيات الذكاء الاصطناعي، وبالتالي فإشراك أصحاب المصلحة أمرٌ مهم، لأنهم يمتلكون حدة نظر تجعلهم يسلطون الضوء على جوانب من التقنيات التي يفتقد إليها عامة النّاس، فقد يقومون بتغيير المفاهيم الأخلاقية بغية ملاءمتها لوظيفة الآلة وقد يلغون بعضها إذا اقتضى الأمر ذلك، نظرا لأنّ بعض المفاهيم الأخلاقية غالبًا ما تؤدي إلى تصميمات مختلفة للتكنولوجيات، مما يضطرهم إلى تقييم المفاهيم الأخلاقية وهذا من شأن مجال الهندسة المفاهيمية. (Stefan Buijsman and others, 2023, page 13) ونحن نكتب في هذا الطرح تبادر إلى أذهاننا هذا التساؤل: هل يمكن صناعة آلة فاضلة تحاكي في فضيلتها الأفعال الإنسانية؟ ففي حالة تعذر ذلك، فإن تكنولوجيا الآلات تبقى محل مشكوكية، لأنها تتحرك وفق نهج أخلاقي تحتي (واقعي) وليس نهجًا أخلاقيًا فوقيًا (مثالي).

نعود إلى مبحثنا حيث ذكرنا مصطلح الهندسة المفاهيمية الذي يدخل في مجال التصميمات التكنولوجية، فشئنا أن نُقرّب المسألة للقارئ، واقترحنا أن نضرب مثالاً بتطبيق الفيسبوك الأكثر استعمالاً، حيث يدخل مفهوم الخصوصية ضمن مجال الهندسة المفاهيمية، الذي يحمي المستهلك من ضرر التطفل ويحفظ حريته في عرض هويته، فأتيح له تقنية قفل الحساب حتى لا يتعرف عليه الآخرون وفي هذا ضمان لخصوصيته التفاضلية.

ومما لا شكّ فيه، أنّ التغيير من القيم الأخلاقية أمرٌ وارد في الأخلاقيات العواقبية، التي تركّز في تعاملاتها على النتائج على حساب الأفعال، وهي تفعل ذلك من أجل تحصيل المنفعة التي حصروها في أبعاد شخصية ومجتمعية وتتمثل في: (Alexander Mitov, 2025, page 5-6)

- صحة، دخل، تعليم، سعادة، السلامة (البعد الشخصي).

- الاقتصاد والتكنولوجيا والبيئة (المجتمعي والعالمي).

بناء على متبيناتنا التي عرضنا فيها حضور النزعة العواقبية في أخلاقيات الذكاء الاصطناعي، من خلال وقوفنا عند أهم الخصائص أو الأحرى الركائز التي تتحرك على هديها الأخلاق العواقبية، والتي سبق ذكرها في متن المبحث (النتائجية، المنفعة، المصلحة)، نتساءل: هل نجحت الأخلاق العواقبية في التقييم الأخلاقي والتسيير الوظيفي المقبول لأخلاقيات الذكاء الاصطناعي؟ بيان ذلك سيكون في المبحث الموالي.

ثالثا: انعكاسات الأخلاق العو اقبية على أخلاقيات الذكاء الاصطناعي

تعرضنا في المبحث السابق إلى نقطة المناسبة، والتي ذكرنا فيها مدى ملائمة الأخلاق العواقبية لتصميمات أخلاقيات الذكاء الاصطناعي، من قبيل أن تلك الأخلاقيات –العواقبية- تتسم بالعملية كونها أخلاقَ محايثة تنطلق من الواقع، وليست أخلاقًا متعاليةً تنطلق من معايير فوقية، الأمر الذي جعلها تحقق نتائج سريعة تضمن رفاهية الإنسان على الصعيد الشخصي ورفاهية المجتمع على الصعيد العام. والسؤال المطروح ههنا: هل حقق حضور النزعة العواقبية في أخلاقيات الذكاء الاصطناعي الرفاهية المرجوة؟

يرى العديد من المفكرين أنّ العواقبية كنهج أخلاقي لا يصلح لتسيير أخلاقيات الذكاء الاصطناعي، حيث سبب حضورها العديد من المشكلات التي خيبت توقعات وكلاء الذكاء الاصطناعي من جهة والمستخدمين المتفاعلين من جهة أخرى. وقد اخترنا أهمها نعرضها كما يلي:

1-عدم دقة التنبؤ: إنّ القيمة الأخلاقية في المتن العواقبي تتحدد بقدرة الفاعل الأخلاقي على التنبؤ بحسن النتائج بصرف النظر عن الفعل المتبع، وهذا يعد سلوكًا هامًا في تصميم الأنظمة الخوارزمية للذكاء الاصطناعي، إلا أنّه في الوقت نفسه يعتبر من بين أصعب المشاكل التي يواجها وكلاء الذكاء الاصطناعي، فيما يخص توقعاتهم بمدى تأثير نظام الذكاء الاصطناعي قبل وبعد استخدامه، وهذا راجع إلى كون هندستنا التكوينية عاجزة على التنبؤ بكل شيء، بسبب التغير في القيم مع مرور الوقت. وقد استدعى هذا الخوف من فشل تحقيق النتائج المرجوة في المستقبل، دعوة الاتحاد الأوروبي القائمين على الذكاء الاصطناعي إلى مراقبتهم المستمرة لتلك الأنظمة الخوارزمية، حذرًا من حدوث الأسوأ (Stefan Buijsman and others, 2023, page 14-15) وهذا دليل على أن التنبؤات التي يضعها الوكلاء لا تتسم بالمطلقية والموثوقية، وقد أدى هذا المشكل إلى قيام مشكلة أعظم المتمثلة في عدم اليقين بما سنكون عليه في المستقبل من جرّاء استخدامنا لهذه الأنظمة المعقدة. بيانها في النقطة الموالية.

2-عدم اليقين: ونقصد هنا انتفاء اليقين بالعواقب الجيدة التي ستتحقق في المستقبل، وبالتالي فالمستقبل يعد هاجزًا غامضًا أمام وكلاء ومصممي استخدامات الذكاء الاصطناعي، فحتى بالرجوع إلى ممارستنا العملية اليومية، نجد أنفسنا عاجزين أمام التوقع المطلق لما ستكون عليه عواقب قرارتنا المنوطة بالنتائج على حساب الأفعال، وكل هذا يجعلنا نعيش موقفًا أخلاقيًا مبنيًا على الاحتمال (المجهول). كما أن قدرة التوقعات تختلف من شخصٍ إلى آخر. وعليه فإنّ المشكلة ليست في عدم اليقين بالعواقب فحسب، بل تمس ذواتنا كذلك فنحن نجهل ما إذا كانت نماذجنا الخوارزمية ستلائم حالة العالم مستقبلا (Dallas Card, Noah Smith, 2023, page 05)، بمعنى أن النهج العواقبي الذي يبني صورة خيالية لما سيكون عليه العالم مستقبلاً في ظل التغيرات التقنية، يجعل من أخلاقيات الذكاء

الاصطناعي في حالة شك دائم وخوف مستمر من المستقبل سواء القريب منه أو البعيد.

ولعل أولى المعيقات التي تعترض صحة توقعاتنا المستقبلية تتمثّل بمعرفتنا الناقصة للبنية السببية للكون، مما يجعل قدرتنا على التنبؤ بالمستقبل محدودة بسبب الافتقار –كذلك- إلى الدقة في معرفتنا بالحاضر (Dallas Card, Noah Smith, 2023, page 05) خاصة وأنّ الفاعل الأخلاقي هنا نظامٌ خوارزميٌ معقد وليس الإنسان المدرك المريد العارف لنتيجة تصرفه في موقف مباغت، حيث تؤدي الإرادة الفاعلة دورَها في تحديد القيمة الأخلاقية، وبالنظر إلى البرمجيات الخوارزمية نجد أنّ نظام الآلة يفتقر إلى أحد شروط الأخلاق التي تتمثّل بإرادة الاختيار – في حين يمثّل العقل الشرط الآخر- صحيح أن الآلة قد تختار التصرّف المبرمج ولكن إرادتها فاعلة مُفَّعلة وليست فاعلة منفعلة. وهنا نطرح التساؤل التالي: هل يمكن للآلة أن تضحي بنفسِها في موقف مفاجئ؟

وبحديثنا عن المواقف المفاجئة، نستحضر عامل الصدفة، الذي يعتبر عائقًا في توقع العواقب التي لم تكن في الحسبان، حيث «تعتمد التبعية على افتراض قدرة الفاعلين الأخلاقيين على تصور عواقب نواياهم وأفعالهم بوضوح وتميز. غير أنّها في مجال الذكاء الاصطناعي، وفيما يخصّ الابتكارات بشكلٍ خاص، لم يتم إثبات هذه الفرضية، على الأقل ليس بشكلٍ مطلق، بالنظر إلى الواقع، لم يتم تنفيذ عدد من الابتكارات بالطريقة التي صُممت بها في البداية، وبمجرد طرحها في السوق، نادرًا ما يتم استخدامها وفقًا لتوصيات مصمميها، حيث يساء استخدامها من قبل المستخدمين. وعليه، فإن النهج التبعي يواجه بطريقة ما تحديًا خطيرًا متمثلاً بحالات الصدفة المميزة لمواقف الابتكار (Ayoub El Azami El Idrissi et autres, 2024, page 07) ومن حالات الصدفة التي يمكن ادراجها ضمن هذا الطرح نذكر – على سبيل المثال لا الحصر- الجريمة الالكترونية، حيث لم يكن في الحسبان أن يساء استخدام تطبيق الفيسبوك، ليؤدي في آخر المطاف إلى معضلات أخلاقية واجتماعية، ألحقت الضرر بالطرف الآخر، سواء أكان ضرر جسديًا أم نفسيًا. كذلك التطفل والأطفال الذين يستخدمون الفيسبوك ومواقع التواصل الاجتماعي دون السن القانوني، وفي هذا تجاوز لخاصية المراقبة الأبوية.

3-غياب القاعدة الأخلاقية المثلى: بالنظر إلى الأخلاق المعيارية –التي تمثل الأخلاق العواقبية قسيماً منها إلى جانب الأخلاق الواجبية وأخلاق الفضيلة- نرى كل من الأخلاق الواجبية وأخلاق الفضيلة يتخذون قاعدة أخلاقية ثابتة يعملون على هديها في معرفة حسن الفعل من قبحه. حيث تعتمد الواجباتية الأخلاقية على الواجب الأخلاقي، وتعتمد أخلاق الفضيلة على الفاعل الفاضل، في حين نجد العواقبية تُعوِّل فقط على النتائج الحسنة المبنية أساسًا على مجموع التوقعات التي تبقى في آخر المطاف احتمالات يمكن أن تنجح أو تفشل، سواء على الصعيد الذاتاني أو الصعيد الموضوعي. وهناك من يرى أنّ وكلاء الذكاء الاصطناعي قد وقعوا في هذا المأزق العواقبي المتمثل بعدم ثبات القيمة الأخلاقية التي تحركها

مجموع التنبؤات الاحتمالية، وعليه يمكن القول أنّ النزعة النفعية عاجزة على توفير أو تحديد السياسة المثلى الكفيلة بضبط التوقعات، في ظل عدم اليقين الذي نعاني منه بشأن العالم (Dallas Card, Noah Smith, 2023, page 05).

وبالتالي سيبقى مشكل تحديد السياسة المثلى لتقييم أنظمة الذكاء الاصطناعي - التي لن تتحقق إلا بالخضوع للقاعدة الأخلاقية الثابتة - من أصعب المشاكل التي يجب على وكلاء الذكاء الاصطناعي ايجادها، خاصة بعد التحديات التي فرضتها وتيرة الاختراعات والابتكارات، ومع ذلك تظل العواقب طويلة الأجل غير متوقعة إلى حد ما. وهذا لا ينفي الجهود الساعية لإيجاد حلول لهذه المعضلة.

ضمن الطرح نفسه، يرى فريق آخر من الباحثين، أنّ العواقبية عامة والنفعية خاصة، تتجاهل الرغبات الإنسانية باتباعها الطريقة الحسابية لتحقيق النتائج الجيدة، وهم –أي الباحثون- يعتقدون أنّ الطبيعة السياسية للتقييم ينبغي أن يكون البحث فيها عن مجموعة واسعة من وجهات النظر حول ما هو مرغوب فيه والتفكير فيه، ليس من أجل حساب نفعي اختزالي... وبدلاً من السعي وراء تحسين سبل الحياة لمن هم في وضع أفضل عليهم توجيه الجهود لمساعدة أولئك الذين هم أسوأ حالا (Dallas Card, Noah Smith, 2023, page 05) ولعل هذا ما نراه متجليًا في فارق التقدم التكنولوجي واستخدام أنظمة الذكاء الاصطناعي بين الدول المتقدمة والدول المتخلفة.

أدّت هذه المشكلة المتمثلة بغياب الضابطة الأخلاقية المثلى التي من خلالها نقيس حسن النتيجة من قبحها، إلى حدوث مشكلة أخرى تتمثل بصعوبة قياس الرفاهية، فإذا انتفت القاعدة تنتفي الغاية ضرورة. بيّنة ذلك في النقطة الموالية.

4-صعوبة قياس الرفاهية:

تعتبر الرفاهية الناتج الأفضل الذي يقيس به العواقبيون نجاح النتائج الفورية أو المستقبلية. غير أن قياسها يعتبر مشكلاً في حد ذاته، من وجهة تشخيص بعض الباحثين الذين رأوا أن أنّ تجميع الرفاهة يتطلب قياس الرفاهة الفردية، ولكن من غير الواضح ما إذا كان من الممكن قياسها بطريقة تسمح بإجراء مقارنات عادلة، على الأقل في ظل التكنولوجيا الحالية. وحتى إذا قصرنا مجموعة الكيانات ذات الصلة الأخلاقية على البشر، فإن قضايا الذاتية، والمزاج، والإبلاغ الذاتي تجعل من الصعب إن لم يكن من المستحيل إجراء مقارنة ذات مغزى بين الأفراد... خاصة وأن الوسط الذي يقاس فيه ذلك هو وسط افتراضي، ينطوي على صعوبات حسابية (Dallas Card, Noah Smith, 2023, page 05) بمعنى أن العامل النفسي يؤثر في قياس مدى تفاعل الفرد مع تطبيقات الذكاء الاصطناعي، وبالتالي فما يطرحه الوكلاء من قيم حسابية للرفاهية تظل محل شك، طالما أن العامل الذاتي (من مزاجيات وعواطف وميولات نفسانية)

تدخل في قياس منسوب الرفاهية. فحتى لو افترضنا اتباع طريقة مثلى في حساب الرفاهية فربما سينجح ذلك في عمود العالم الخارجي، غير أنّ ذلك سيكون أصعب بكثير من حسابه في العالم الافتراضي، لكون هذا الأخير لا يتسم بالشفافية والمصداقية المطلقة.

إلى جانب العامل الذاتاني نجد العامل الاجتماعي الذي يترشح منه مجموع القيم والثقافات التي من شأنها أن تساهم في بناء حياة الفرد، وهذا الأخير يعتبر عائقًا أساسيًا أمام قياس الرفاهية لأنّ تحديد القيم الأخلاقية الجيدة أو السيئة يتقلب باختلاف الثقافات. وهذا ما نسميه النسبانية الأخلاقية، وهو المفهوم الأكثر قبولاً بين السكان بشكلٍ عام. ومن ناحية أخرى، نرى أنّ النّاس لديهم حدسٌ واقعي، أي أنهم يعتقدون أنّ هناك أشياءَ خاطئة حقًا، مهما كانت الثقافة. والمثال الذي يُضرب به كثيرًا هو مثال العبودية التي كانت مقبولة في وقت من الأوقات، والتي لم تعد كذلك (Pascal Hérard, 2021) بمعنى أنّه ليس بالضرورة أن تكون النتيجة –مثلا- التي يسعى إلى تحقيقها وكلاء الذكاء الاصطناعي خادمة في الوقت نفسه للمنظومات الليبرالية والمنظومات الدينية، بحكم اختلاف الثقافات العرفية وكذلك التوجهات المعتقدية. إضافة إلى الوازع الثقافي والديني، فإن أي واحد قد يحدس مباشرة أنّ ذلك الشيء قد يلحق الضرر، فنتجنب استخدامه، أو نتوخى الحذر عند استخدامه. وكل ذلك من شأنه أن يؤثر في حساب نسب الرفاهية العالمية المترتبة عن أخلاقيات الذكاء الاصطناعي.

ومن القيم الأخلاقية التي نوقشت لوضع أساسٍ متين لأخلاقيات الذكاء الاصطناعي، العدالة، غير أنّه وبالرجوع إلى الدراسات الأخلاقية الأولى نرى أنّ هذه القيمة ذات دلالة زئبقية، بمعنى أن مفهومها غير محدد، بحكم تداخله تارة بالثقافات الشعبية وتارة أخرى بالمعتقدات الدينية، وهذا ما أحدث مشكلاً في برجمة أنظمة الذكاء الاصطناعي، منها التي تخدم أصحاب المصلحة، ومنها التي تميل إلى الجنس الأبيض على حساب الجنس الأسود، ومنها التي تخدم أمة دون غيرها. (Dallas Card, Noah Smith, 2023, pages 3-9-10) (Stefan Buijsman and others, 2023, 47-48) على سبيل المثال لا الحصر، لماذا حين تنشر منشورات في حساب الفيسبوك –كنموذج- حول القضية الفلسطينية والتي لها مساس بالاحتلال الصهيوني، يحظر المنشور ويعاقب صاحب الحساب، على غرار القضايا الأخرى التي تحكي قضايا سياسية أخرى؟ والسؤال المطروح ههنا: هل تؤدي مثل هذه الحسابات الأخلاقية المبنية على عدالة نسبية إلى قياس مضبوط لنسب الرفاهية الناتجة عن استخدام أنظمة الذكاء الاصطناعي؟

5-العامل الزمني: يعتبر الزمن عاملاً مهما في حساب العواقب سواء القريبة منها أو البعيدة، المترتبة عن أنظمة الذكاء الاصطناعي، وقد شَكَّلَ هذا الأخير عائقًا، ولعل أولى العوائق تلك المتمثلة بتخيل العالم بناء على مجموع احتمالات انطلاقًا من الحاضر، خاصة وأنّ الأمر فيه مساسٌ بالجانب القيمي لحياة الفرد. «حيث يتم بناء مستقبل استنادًا إلى لائحة عامة لحماية البيانات، والتي قد تخلف تأثيرات هائلة، لأنّها

ستحدد طريقة تفكيرنا حول القيم (خاصة قيمة العدالة) في المستقبل، وكذلك المعيار الذي نقيس به معايير الحسن والقبح، مما يفرز إشكالاً آخرَ يتعلق بمصداقية التنبؤ (Dallas Card, Noah Smith, 2023, pages 08).

وسياقًا لذلك المحكي، دعا أحد المتدخلين الباحثين بضرورة إعطاء أهمية لأولئك الذين سيعيشون في المستقبل مثلما أعطيناها للأشخاص الذين يعيشون في الحاضر (Dallas Card, Noah Smith, 2023, page 05) وتبعًا لذلك نطرح التساؤل التالي: هل من الحق أن نحدد طريقة عيش الأجيال اللاحقة؟ وهل سيرضون بذلك؟

أثارت تلك المشكلة –المتعلقة بالأجيال اللاحقة- إشكالية المجتمع المرغوب فيه في المستقبل، حيث تتحمل النفعانية الأخلاقية –المعتمدة في تقييم نتائج الذكاء الاصطناعي- مسؤولية التفكير في الأجيال القادمة مقارنة بالأجيال الحالية، نظرًا لأنّ أولئك الذين لم يولدوا بعد غير قادرين على الدفاع عن مصالحهم بشكل مباشر (Dallas Card, Noah Smith,2023, pages 05) وهنا أُشتهت النفعانية كنهج للأنظمة الخوارزمية في كونها ركّزت على الجانب الاجتماعي في سبيل السعي لتوفير حياة على قدر عالي من الرفاهية، وذلك على حساب الجانب القيمي، حيث لم يفكر في ردّة فعل الأجيال اللاحقة من أفق أخلاقي للعالم الذي حُتم عليهم العيش فيه دون إرادة منهم. وهنا نرى إسقاط لأحد شروط الأخلاق المتمثّل بحرية الاختيار. ونحن بدورنا كباحثين نثير التساؤل الآتي: هل من الجائز أخلاقيًا أن يصبح مصير الأجيال اللاحقة نظامًا من الأنظمة الخوارزمية؟

6-الأمن والسلامة: اضطررنا إلى تأخير هذه النقطة المهمة التي تؤخذ عليها العواقبية المسيّرة لأنظمة الذكاء الاصطناعي، فبعد عرضنا للسلبيات هذا النهج الأخلاقي والتي ضمّناها في النقاط السالف ذكرها، هل يُتوقع منه أن يكون حافظًا لخصوصية المستخدمين لتطبيقات الذكاء الاصطناعي؟ هناك من يرى أنّ بالرغم من إقرار الاتحاد الأوروبي بضرورة قانون يضمن خصوصية المستخدمين من التطفل والاختراق والمعروف بكود: GDPR أشد قوانين الخصوصية والأمن صرامة في العالم، وهو عبارة عن لائحة عامة لحماية البيانات، إلا أنّ الاستخدام غير الأخلاقي ومعالجة البيانات من خلال استخدام الذكاء الاصطناعي لا يزال من الممكن أن يعاقب عليه القوانين المعمول بها (Alexander Mitov, 2025) ولعل سرقة البنوك واختراق الحسابات أبرز مثال يمكن إدراجه ضمن هذا الطرح، والذي يوضح ضعف أنظمة الحماية.

وتُعتبر شركة كامبريدج أنالتيكا Cambridge Analytica خير مثال على التجاوز الأخلاقي واختراق البيانات الشخصية، حيث أَقدمت هذه الأخيرة على جمع بيانات ما يقارب 50 مليون شخص من مستخدمي الفيسبوك في الترويج للانتخابات السياسية. مما دفع إلى المطالبة بتحديث نظام جديد يمكن الوثوق به في حماية خصوصيات المستهلكين. إنّ هذه الحادثة التي زعزعت ثقة الأشخاص الذين يستخدمون منصات

التواصل الاجتماعي (Alexander Mitov, 2025) ولعل السعي وراء تحقيق أفضل النتائج هو الذي دفع بقادة تلك الشركة إلى ارتكاب ذلك الخطأ الشنيع، الذي ينتهك خصوصيات الأفراد. ويمكننا ضرب مثال آخر بالحادثة التي اعتبرها قادة حزب الله خرق أمني، استهدف الهواتف النقالة التي استوردوها، مما أدى إلى انفجارها دفعة واحدة، كانت الحادثة قد وقعت في لبنان في 17 سبتمبر 2024 (wikipedia, 2025) في حين أرجع البعض الآخر سبب الانفجار إلى خلل تقني وليس نظامي. تتعدد الأمثلة غير أنّ كلها تصب في إشكالية واحدة تتعلق بالأمن والخصوصية من جرّاء استخدام أنظمة الذكاء الاصطناعي.

عملاً بتلك المتبيّنات، التي تستوضح مساوئ اتباع النهج العواقبي في تسيير أنظمة الذكاء الاصطناعي، والذي تتحدد فيه القيمة الأخلاقية بحسن النتيجة، حيث خلّف ذلك عواقب وخيمة خيبت التوقعات المرجوة، فألحقت بذلك الضرر بمستخدمي تلك الأنظمة، كان أهمها تلك المتعلقة بالحتمية المصيرية للأجيال اللاحقة من جهة، والأمن والخصوصية من جهة أخرى.

خاتمة

بعد العرض البسيط الذي وضحنا فيه حضور العواقبية الأخلاقية كنهج متبع في تسيير أنظمة الذكاء الاصطناعي، والذي عرضنا فيه صلاحية النهج من أفق عدة نقاط من جهة، والانعكاسات السلبية لذلك النهج من جهة أخرى، وكذلك حصرناها في نقاط أسلفنا ذكرها. وشئنا ضمن هذا الختام، أنّ نجري مقاربة –دون التطويل الممل والتقصير المخل- بين المبنى العواقبي والمباني المعيارية الأخرى المتمثلة بالواجبية الأخلاقية وأخلاق الفضيلة كبدائل أخلاقية مرشدة للذكاء الاصطناعي.

تتحدد القيمة الأخلاقية في متن العواقبية –كما أسلفنا الذكر في العرض- بتحقيق أفضل النتائج، وقد ترتب على ذلك العديد من المعضلات الأخلاقية التي لم تكن في الحسبان. فإذا كانت القيمة الأخلاقية المنوطة بالنتيجة قد تخطئ فلماذا لا نركز على الفعل الذي من خلاله تتحدد النتيجة بصفة معلومة وغير احتمالية –كما هو الحال مع العواقبيين- وهذا ما تعمل على هديه أخلاقيات الواجب.

وضمن هذا الأفق الذي يفتح السبل للتقسيمات الأخلاقية في أن تكون لها فعالية في توجيه الأنظمة الخوارزمية بدلاً من الانغلاق على النموذج التبعي، يرى بعض المفكرين أنّه قد بات من الضروري أن نستدعي كانط من جديد، من أجل صياغة التشريع الأخلاقي الذي يحدد القدرة البشرية على صياغة حكم أخلاقي مستقل، وهذا يعني الحكم على أساس القدرة في توجيه تصرفات الفرد من خلال التفكير النزيه، بدلاً من التفكير المبني للمجهول القائم على مجموع احتمالات غير متقين من حدوثها على النحو المطلوب. أما بخصوص أخلاقيات الفضيلة القائمة على الممارسة المستمرة للفضائل (Ayoub ElAzami El Idrissi et autres, 2024, page 08) فيمكن أن تساعد وكلاء الذكاء الاصطناعي في تركيز الاهتمام على الفاعل الأخلاقي، بدلاً من السعي وراء تحقيق أفضل النتائج لتوفير الرفاهية. سبق الإشارة في متن البحث إلى

مفهوم العدالة من منظور نفعي، حيث اعتقد بعض المفكرون أن النفعية كمعيار لفضيلة العدالة غير صالح، حيث التمس فيه بعض المستخدمين لأنظمة الذكاء الاصطناعي نوعًا من العنصرية العرفية والعرقية، فالوضع سيكون أفضل لو تم الأخذ بالمفهوم الأفلاطوني للعدالة.

إلى جانب النسقين الأخلاقيين –الواجبي والإريتاني (الفضيلة)- يمكن أن تكون كذلك الأخلاق الدينية بديلاً عن النسق الأخلاقي العواقبي، حيث تتحدد القيمة الأخلاقية ضمن هذا القسم الأخلاقي بمدى العمل بالتكاليف الإلهية، التي هي تكاليف فوقية منزهة، ولا قيمة تعلو على الإرادة الإلهية – لا الفعل المقيد بالواجب ولا الفاعل الأخلاقي ذا التكوين الناقص- والمتمعن هنا -مع التحفظ برأيي الشخصي- يرى أن جميع القيم التي اختلفت بين الأنساق الأخلاقية اجتمعت وتفاعلت مع بعضها ضمن هذه الأخلاقية ذات البعد الإلهي التوحيدي. فبمقتضى الأمر بالمعروف (النسب الارسالية) والنهي عن المنكر (النسب الامساكية) يكون الفعل حسنا متى امتثل لشرع الله وقبيحًا متى اعترض وتمرد على ذلك، بحيث تكون النتيجة يقينية، طال أمدها أم قَصُرَ، وسيترتب بفعل الممارسة المستمرة للفضيلة فاعلاً فاضلاً، مما يورثه السعادة الدنيوية والأخروية. ولنفترض أن هذه الأخلاقية معيارٌ ثابتٌ يمكن العمل على هديه في تحديد مسار أخلاقيات الذكاء الاصطناعي بحيث تتضح نتائجه بشكل واضح بعيدًا عن التكهنات والتوقعات المبنية على احتمالات حسابية يمكن أن تصيب ويمكن أن تفشل. بمعنى أن الأخلاقيات العواقبية تتحرك في حدود الإمكان، وهذا ما جعل أساسها هش عرضة للنقد والشك.

قائمة المراجع

1) مونيك كانتو سبيربير، روفين أدجيان، (2004)، الفلسفة الأخلاقية، ترجمة جورج زيناتي، ط1، لبنان، دار الكتاب الجديد.

2) Emmanuel Kant, (1888), critique de la raison pratique, nouvelle traduction française, avec un avant-propos, sur- la philosophie de Kant en frange, de 1773 a 1814, des notes philologiques et philosophiques par F. Picavet, agrégé de philosophie, paris, ancienne librairie germer ballerai Félix Algan, éditeur ν ; 1 108, boulevard Saint – Germain, 408.

3) Josiane Boulad-Ayoub, (1990), Fiches pour l'étude de Kant, Québec, Recherches et Théories. Collection: Symbolique et idéologie.

4) Konstantin Büchler, Qu'est-ce que l'éthique des vertus?, (2014), HAL Id : hal-00994574 https://hal.science/hal-00994574 Preprintsubmitted on 1 Jun 2014.

5) David Massey, virtue ethics, PHI 105: Introduction to Ethics, India.

6) Le bonheur, http://www.guillaumenicaise.com.

7) Jeremy Bentham, (2000), An Introduction to the Principles of Morals and Legislation, Kitchener, Batoche Books.

8) Jean-Pierre Cléro, (2011), L'utilitarisme est-il une éthique acceptable ? Revue d'études benthamiennes, N°9, Pages.(484-464)

9) Josiah Della Foresta, (2020). Consequentialism & Machine Ethics: Towards a Foundational Machine Ethic to Ensure the Right Action of Artificial Moral Agents. Montreal AI Ethics Institute, Pages .(14-01)

10) Dallas Card, Noah. Smith,(2020), On Consequentialism and Fairness, frontiersin in Artificial Intelligence | www.frontiersin.org , Volume 3, Pages.(11-01)

11) Stefan Buijsmanand others, (2023), Ethics of Artificial Intelligence. In Nathalie Smuha, Cambridge Handbook on the Law, Ethics and Policy of AI, Pages.(17-01)

12) Ayoub El Azami El Idrissi et autres, (2024), Nouvelles technologies digitales et risques éthiques pour les ingénieurs : limites des cadres existants et réflexion dans le cadre de la performance industrielle, INSA Hauts, France .

13) Allison Marchildon et autres, (2023), L'éthique au cœur de l'IA, Obvia, Québec, ISBN : 978-2-925138-26-6 DOI : 10.61737/MDHP6080. Pages.(24-04)

14) Schefczyk, Michael (2011). Mill's Ethics. In James Fieser & Bradley Dowden, Internet Encyclopedia of Philosophy.| ISSN 2161-0002. https://iep.utm.edu/home/about .

15) Alexander Mitov, (2021), Ethical use of artificial intelligence through the Utilitarianism perspective, https://essay.utwente.nl/86726/. 10/02/2025. 12.29 .

16) Pascal Hérard, Quelle éthique pour l'intelligence artificielle ? (2021), https://information.tv5monde.com, Le 20 déc. 2019 à 11h00 (TU) Mis à jour le 24 déc. 2021 à 15h01 (TU), 28/05/2024. 20.32.

17) https://ar.wikipedia.org/ 21/04/2025. 18.45.