المجلة الأردنية في القانون والعلوم السياسية

اسم المقال: الذكاء الاصطناعي بين الاعتبار الأخلاقي والمقياس القانوني "أين نحن؟" دراسة نقدية في الفلسفة والتأصيل
اسم الكاتب: أ.د. محمد عرفان الخطيب
رابط ثابت: https://political-encyclopedia.org/index.php/library/8224
تاريخ الاسترداد: 2026/11/02 07:36 +03

الموسوعة السياسيّة هي مبادرة أكاديمية غير هادفة للربح، تساعد الباحثين والطلاب على الوصول واستخدام وبناء مجموعات أوسع من المحتوى العلمي العربي في مجال علم السياسة واستخدامها في الأرشيف الرقمي الموثوق به لإغناء المحتوى العربي على الإنترنت.
لمزيد من المعلومات حول الموسوعة السياسيّة – Encyclopedia Political، يرجى التواصل على info@political-encyclopedia.org

استخدامكم لأرشيف مكتبة الموسوعة السياسيّة – Encyclopedia Political يعني موافقتك على شروط وأحكام الاستخدام المتاحة على الموقع https://political-encyclopedia.org/terms-of-use

تم الحصول على هذا المقال من موقع المجلة الأردنية في القانون والعلوم السياسية جامعة مؤتة ورفده في مكتبة الموسوعة السياسيّة مستوفياً شروط حقوق الملكية الفكريّة ومتطلبات رخصة المشاع الإبداعي التي ينضوي المقال تحتها.

Artificial intelligence between ethical consideration and legal standard "Where are we?" A critical study in philosophy and rooting

Ph.D. Mhd Arfan Alkhatib

Prof. in civil law – Ahmad bin Mohammad military College

Doha – Qatar

Received: 02/04/2023

Revised: 02/07/2023

Accepted: 03/07/2023

Published: 30/09/2023

DOI:

https://doi.org/10.35682/jjlps.v15i3.572

Corresponding author:

Alkhatibur@abmmc.edu.qa

All Rights Resaved for Mutah University, Karak, Jordan.

All Rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means: electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the publisher.

Abstract

This paper aims to provide a critical and philosophical study of one of the most complex aspects of the ethics of artificial intelligence and hopefully resolve the question of where our location concerning the ethics of artificial intelligence is. The question is examined from two viewpoints, the ethical and the legal. In the first section, the study attempts to define the expected artificial moral behavior, determine its nature, and how it can be achieved. In the second, the study attempts to clarify the measurement of the ethical dimension of the expected artificial behavior and whether we are socially or legally equipped to deal with ethical artificial intelligence.

The research concludes that ensuring that artificial intelligence behaves ethically must take precedence over the pure concept of the ethics of this intelligence. Moreover, the technical standards of this intelligence must be subordinated to human standards. At the same time, the study discovers a coherent knowledge circle, i.e., the study of the morality of artificial intelligence leads to enhanced artificial intelligence. In contrast, the study of artificial intelligence leads to a better understanding of the morality of artificial intelligence. The study recommends that each be studied to understand the other better.

Keywords: Ethical AI. Artificial social conscience. Legal Standard of AI Ethics. Ethics of AI. Moral AI. formative concept of AI

.

الذكاء الاصطناعي بين الاعتبار الأخلاقي والمقياس القانوني "أين نحن؟"

دراسة نقدية في الفلسفة والتأصيل

أ. د. محمد عرفان الخطيب

أستاذ القانون المدني

قسم القانون–كلية أحمد بن محمد العسكرية

قطر – الدوحة

تاريخ الاستلام: 2023/04/02
تاريخ المراجعة: 2023/07/02
تاريخ موافقة النشر:2023/07/03
تاريخ النشر:2023/09/30

الباحث المراسل:

Alkhatibur@abmmc.edu.qa

حقوق النشر محفوظة لجامعة مؤتة، الكرك، الأردن.

جميع الحقوق محفوظة، فلا يسمح بإعادة طباعة هذه المادة أو النقل منها أو تخزينها، سواء أكان ذلك عن طريق النسخ، أم التصوير، أم التسجيل، أم غيره، وبأية وسيلة كانت: إلكترونية، أو ميكانيكية، إلا بإذن خطي من الناشر نفسه.

الملخص

عبر دراسة نقدية في الفلسفة والتأصيل، يتناول موضوع البحث أحد جوانب أخلقة الذكاء الاصطناعي الأكثر تعقيداً، المرتبطة بتحديد: "أين نحن؟" من أخلاقيات الذكاء الاصطناعي. وذلك ضمن سؤالين قوامها الاعتبار الأخلاقي والمقياس القانوني. عارضاً لذلك وفق مخطط بحثي، حاول في قسمه الأول تحديد ماهية السلوك الأخلاقي الاصطناعي المُنتظر، لتحديد كُنه هذا السلوك، وكيفية الوصول إليه. بينما حاول قسمه الثاني، توضيح مقياس البعد الأخلاقي للسلوك الاصطناعي المُنتظر، وهل نحن أمام ذكاء اصطناعي أخلاقي بمقياس اجتماعي أم قانوني.

وقد خلص البحث إلى نتيجة معرفية مفادها أن ضمان تصرف الذكاء الاصطناعي بأخلاقية، يجب أن يكون مقدماً على المفهوم الصرف لأخلاق هذا الذكاء. وبأنه يجب تحييد المعايير التقنية ضمن هذا الذكاء لمصلحة المعايير الإنسانية فيه. بذات الوقت، توصل البحث إلى حلقة معرفية مترابطة قوامها: أن أخلقة الذكاء الاصطناعي هي مفتاح تهذيب الذكاء الاصطناعي نفسه، وبأن تهذيب الذكاء الاصطناعي هي البوابة الموصلة لأخلقة الذكاء الاصطناعي. مُوصياً بضرورة احترام هذه المعادلة في أخلقة أخلاقيات الذكاء الاصطناعي، بوصفها الدائرة المغلقة لحلقتين متكاملتين كلٌ منهما: البوابة والمَخرج للأخرى.

الكلمات المفتاحية: الذكاء الاصطناعي الأخلاقي. الضمير الاجتماعي "الاصطناعي". المقياس القانوني لأخلاقيات الذكاء الاصطناعي. أخلقة الذكاء الاصطناعي. المفهوم التكويني للذكاء الاصطناعي.

"المخطط"

مقدمة

المطلب الأول

ماهية السلوك الأخلاقي الاصطناعي المُنتظر

الفرع الأول

كُنه السلوك الأخلاقي الاصطناعي المُنتظر

أولاً–ذكاء اصطناعي أخلاقي أم يتصرف بأخلاقية؟!

ثانياً–ذكاء يتصرف بإدراك رياضي أم أخلاقي؟!

الفرع الثاني

كيفية تحقق السلوك الأخلاقي الاصطناعي المُنتظر

أولاً–تعزيز أخلاقيات مطوري هذه البرمجيات

ثانياً–تعزيز دائرة التشاركية في المفهوم التكويني للذكاء الاصطناعي

المطلب الثاني

مقياس البعد الأخلاقي للسلوك الاصطناعي المُنتظر

الفرع الأول

المعيار الأحادي والصورة المُنتقصة

أولاً–الضمير الاجتماعي "الاصطناعي" ... المقياس المُفتقد

ثانياً–الالتزام القانوني الصرف ... المقياس المُنتقد

الفرع الثاني

التآزر الأخلاقي والقانوني والصورة الكاملة

أولاً–التآزر ... الحل التأصيلي الأمثل

ثانياً–التآزر ... الحل الاستشرافي الأمثل

الخاتمة

أولاً– أولوية التصرف بأخلاقية على التصرف الأخلاقي

ثانياً– تحييد المعايير التقنية لمصلحة المعايير الإنسانية

مقدمة

تتعدد المفاهيم الفلسفية للأخلاق بين مفاهيم عدة، منها المفهوم الشخصي المرتبط بشخص الفرد والمفهوم المجتمعي المرتبط بروح الجماعة التي ينتمي إليها الفرد؛ ومنها المفهوم الوضعي المرتبط بما هو قائم من هذه الأخلاق والمفهوم الميتافيزيقي لها المرتبط بما وراء الواقع ضمن ما يعرف بالأخلاق الفاضلة ... وغير ذلك من مفاهيم (Samuel & Henri, 2000, p. 5). رغم ذلك، وبعيداً عما سبق، ودون الخوض في التعريفات الفلسفية المعمقة للأخلاق (Ricoeur, 1988, pp. 42-45) (Meirieu, 1992)(. Samuel .Britannica, 2007 et Henri 2000,) ، لعل تعريف الفقيه الفرنسي "J-J. Rousseau" للأخلاق بكونها: "مجموعة السلوكيات المتعلقة بالخير والشر المُستقرة في الجماعة الإنسانية"(Rousseau, 1996, P. 329). هو الأقرب للبساطة وإلى الارتباط بموضوع البحث (Groupe d'experts indépendants de haut niveau sur l'intelligence artificielle, 2019. Ethics Guidelines for Trustworthy, 2020.) ، باعتباره يُعبر عن السلوكيات المتعلقة بهذين المفهومين (الخير والشر) المستقرين في ضمير الجماعة الإنسانية*.

غير أن هذا التعريف ورغم بساطته اللغوية ووضوحه الفلسفي، حينما يُسقط على الذكاء الاصطناعي، فإنه يطرح إشكاليات متعددة في تحديد كيفية الإقرار بالبعد الأخلاقي لهذا الذكاء، تحديداً ضمن معضلتين رئيسيتين: الأولى، كيف يمكن لنا اعتبار أن الذكاء الاصطناعي ذكاءً أخلاقياً؟! والثانية، ما هو المقياس أو المؤشر المعياري لهذه الأخلاق؟ إذ يجب ألا ننسى أننا في مختلف هذه الفرضيات من الخوارزميات، إنما نُدخل الرقمي في الإنساني، محاولين تحويل هذه الخوارزميات لمُستشعرات حِسية لصحوة الضمير أو غفلته!

من هنا تأتي فكرة هذه الورقة البحثية التي تعرض لجَدلية الاعتبار الأخلاقي في هذه الأخلاقيات والمقياس القانوني الواجب تبنيه حيالها، وفق نقاش يرتبط بالأصل الفلسفي لأخلاقيات هذا الذكاء، بما فيه البعد الأخلاقي المطلوب فيه ومنه. محاولةً، وفق منهج تحليلي مُعمق، يتناول الموضوع من الناحيتين الفلسفية والتأصيلية، وتكون الإجابة عن هذين السؤالين من خلال معرفة أين نحن من هذين السؤالين. مُرتكزةً في ذلك إلى العديد من الوثائق الأوربية والغربية†، إضافةً إلى العديد من التقارير والدراسات الأكاديمية ذات

---

*. وفق منظور مجموعة الخبراء المكلفة من قبل المفوضية الأوروبية بصياغة القواعد الأخلاقية للذكاء الاصطناعي الجدير بالثقة، "تشمل الأخلاقيات النظرية أربعة مجالات رئيسية هي: ... وتعتبر أن المجال الرابع المتعلق بالأخلاق التطبيقية هو الأقرب للتطبيق على الذكاء الاصطناعي، كونه يوضح تعامل الأخلاق مع مواقف الحياة الواقعية، من حيث ضرورة اتخاذ القرارات العقلانية ضمن أطر زمنية محددة. ما انعكس في رؤيتها لهذا الذكاء الذكاء، حيث ترى بأن "الأخلاق هي مجال فرعي من الفلسفة يتعامل مع أسئلة مثل "ما هو العمل الصالح؟"، "ما هي قيمة الحياة البشرية"، "ما هي الحياة الجيدة؟".

†. من أهمها قرار البرلمان الأوروبي تاريخ 20 أكتوبر 2020 المتعلق بالتوصيات إلى المفوضية بشأن إطار عمل للجوانب الأخلاقية للذكاء الاصطناعي والروبوتات والتقنيات ذات الصلة (2012/2020)، والذي يبين فيه المشرع الأوربي الخطوط العامة لهذا التوجه التشريعي المؤسس لأخلقة الذكاء الاصطناعي وصولا لحوكمته، وفق أربعة مبادئ هي على التوالي: احترام استقلالية الإنسان، ومنع أي

الصلة (Ananya, 2016. Pp. 93-117. Michael & Susan Leigh., 2007, Pp. 15-26. Wagner, 2018, Pp. 84-89. Bostrom, & Yudkowsky, 2014, p. 316, p. 334. Floridi. 2018 A 376. Gotterbarn, Brinkman, and others 2018. Pp. 1-28. Hagendorff, 2020. Pp. 99-120. Larsson, 2020., 23. Morley, Floridi, and others 2019. Pp. 1-21. Pekka, Bauer, and others, 2018. Larsson, 2020, Pp. 437-451. Zhang & Dafoe, 2020. Pp.187-193. Godefroy, 2020. P. 231. Forest, 2017, p. 38s. Pellegrini. 2018, Pp.56-59. Jobin, Marcello, & Effy V. 2019. Pp. 389-99. Larsson, 2020, Pp. 437-451. The European Commission's high-level expert group on artificial intelligence: Ethics guidelines for trustworthy AI. Pp. 1-37. The ethics of artificial intelligence: Issues and initiatives. STOA. PE 634.452. 2020. Study Report | الخطيب، 2021، المقال الأول. 2020، المقال الرابع. 2018. ص. ص. 19–58.).

آملين أن نوفق في تقديم مساهمة علمية بسيطة في هذا المجال، تسلط الضوء على هذه الحيثية، كما تمنح الفقه العربي إضاءة –ولو متواضعة– حول هذا الموضوع الذي كثر الحديث عنه في الأوساط الأكاديمية الغربية، ولا يزال –وفقاً لعلمنا الشخصي– في بداياته في الأوساط الأكاديمية العربية.

مع التأكيد إلى أن طبيعة البحث الفلسفية والجدلية، تُملي أن تكون مناقشة الكثير من نقاطه بطريقة استفسارية مبينة على طرح الأسئلة ومناقشة الفرضيات التي لا ندعي بالضرورة، أننا نمتلك الإجابات القاطعة عليها أو الحلول الواضحة لها. وهو أمرٌ باعتقادنا لا يضير البحث أو يطعن في جودته وأصالته بقدر ما يبين كل منهما. فالطبيعة الجدلية للبحث وموضوعه، الذي هو بالأصل فلسفي لم تستطع، حتى حينه الدراسات الدائرة حوله من حسم الإجابة عنه (Refer to footnote No. 6 of this research)، تجعل من هذه

ضرر، والإنصاف وأخيراً إمكانية التفسير. وقد ترجمت هذه المبادئ الأربعة في في سبعة متطلبات أساسية لذكاء اصطناعي جدير بالثقة وفق مفهوم: العمل البشري والتحكم البشري، والقوة التقنية والأمن، واحترام الخصوصية وإدارة البيانات، والشفافية، والتنوع وعدم التمييز، والإنصاف، والرفاهية الاجتماعية والبيئية، والمسؤولية. كذلك من بين هذه المبادرات، الاقتراح المقدم من البرلمان والمجلس الأوروبي المؤرخ في 21 أبريل 2021 موضوع وضع قواعد منسقة بشأن الذكاء الاصطناعي. "تشريعات الذكاء الاصطناعي وتعديل بعض القوانين التشريعية للاتحاد (COM / 2021/206) النهائي". ونظيره المقدم في ذات التاريخ والمقدم كذلك من البرلمان والمجلس بخصوص القواعد المتعلقة باستخدمات الذكاء الاصطناعي على الآلات والمنتجات ذات الصلة (COM / 2021/202) النهائي. إضافةً لقرار البرلمان الأوروبي بتاريخ 20 أكتوبر 2020 المتعلق بتوصيات المفوضية الأوربية بشأن نظام المسؤولية المدنية للذكاء الاصطناعي (INL) TA-9-2020-0276_EN)2014/2020)). إضافةً (INL) الكتاب الأبيض حول الذكاء الاصطناعي 2020. وكذلك، دراسة CoE بتاريخ 17 ديسمبر 2020 حول جدوى إطار قانوني بشأن الذكاء الاصطناعي(23 (2020) CAHAI) وقرار البرلمان الأوروبي بتاريخ 20 أكتوبر 2020 بشأن حقوق الملكية الفكرية لتطوير التقنيات المتعلقة بالذكاء الاصطناعي (INI)2015/2020). وإشعار مركز التميز بتاريخ 21 نوفمبر 2019 بشأن مسودة التوصية بشأن تأثيرات الأنظمة الخوارزمية على حقوق الإنسان T-PD) (09 (2019). وتقرير مركز التميز بتاريخ 25 يناير 2019 الذكاء الاصطناعي وحماية البيانات: التحديات والحلول الممكنة T-PD) (09Rev (2018). وإرشادات مجلس أوروبا بتاريخ 25 يناير 2019 على الذكاء الاصطناعي وحماية البيانات الشخصية. وقرار البرلمان الأوروبي بتاريخ 12 فبراير 2019 بشأن السياسة الصناعية الأوروبية الشاملة بشأن الذكاء الاصطناعي والروبوتات P8_TA) 0081 (2019). وقواعد القانون المدني على الروبوتات اعتمد البرلمان الأوروبي في 16 فبراير 2017.

الطريقة الفضلى في بحث ومناقشة وعرض الكثير من النقاط المرتبطة به. إذ لا يجب أن يغيب عن ذهنا، أننا نناقش فرضيات وطروحات لا حقائق ومسلمات.

عليه، وضمن ما سبق، سيعرض البحث لمختلف الطروحات السابقة وفق مخطط بحثي، يحاول في قسمه الأول، تحديد ماهية السلوك الأخلاقي الاصطناعي المُنتظر من هذا الذكاء الاصطناعي، ما سيفتح أمامنا فرضيتين فرعيتين: الأولى ترتبط بتحديد كنه السلوك الأخلاقي الاصطناعي الذي نريده من هذا الذكاء ولهذا الذكاء. والثانية تتعلق بتحديد كيفية تحقق هذا السلوك الأخلاقي الاصطناعي المُنتظر. بينما يحاول القسم الثاني، توضيح المقياس أو المؤشر المعياري لهذه الأخلاق، بين الاعتبارين الاجتماعي أو القانوني، مُبيناً عدم صحة الاعتماد الحصري على أي منهما، بذات الوقت الذي يُشدد على صوابية الاعتماد على كليهما، لتغدو النظرة الفردية لأي منهما مُنتقصة، والنظرة التكاملية لك منهما، مُكتملة.

المطلب الأول

ماهية السلوك الأخلاقي الاصطناعي المُنتظر

إن كان السؤال التكيفي البسيط المتعلق بتحديد هل المقصود في أخلاقيات الذكاء الاصطناعي أن نكون أمام ذكاء أخلاقي أم ذكاء يتصرف بأخلاقية؟! وهل يجب أن يكون هذا الذكاء أخلاقياً؟! أم يجب عليه أن يتصرف بشكل أخلاقي ليكون أخلاقياً؟! يطرح مُعضِلةً أخلاقيةً وفلسفيةً كثيراً ما يحتدم النقاش حولها. فإن ما وراء هذا السؤال، مُعضلة فكرية وفلسفية، هي باعتقادنا أكثر عمقاً، تتمثل في تحديد كيف يمكن لنا – نحن البشر – التأكد من وصولنا إلى الذكاء الاصطناعي الأخلاقي الذي نريد ونرغب!

هاتان المعضلتان، يُطرح من خلالهما فرضيتان مفصليتان: الأولى، تتناول تحديد كُنه السلوك الأخلاقي الذي نريده من هذا الذكاء؟! والثانية، ترتبط بكيف يمكننا التأكد من وصولنا إلى الذكاء الأخلاقي الذي نريد؟! سؤالان، بقدر إحساسنا الأولي بسهولة الإجابة عنهما، بقدر يقيننا النهائي بصعوبة ذلك!

الفرع الأول

كُنه السلوك الأخلاقي الاصطناعي المُنتظر

هل الأخلاق هي المؤشر المعياري الصحيح للسلوك الأخلاقي الصحيح، وهل حينما يكون الذكاء الاصطناعي أخلاقياً، يعني ذلك بالضرورة أنه يتصرف بأخلاقية؟! لا نعتقد ذلك! فما بين المعيار والسلوك المُترجم لهذا المعيار، تكمن الفجوة! فالأخلاق –رغم أهميتها– ليست المُعبر عن التصرف الأخلاقي الصحيح دوماً.

بالمقابل، وعلى افتراض قيام الذكاء الاصطناعي بسد هذه الفجوة، والقدرة على الانتقال من المفهوم الأخلاقي إلى التصرف بأخلاقية وفق الشكل الصحيح، فهل يُدرك هذا الذكاء الاصطناعي بوعي من ذاته قيامه بذلك؟ بمعنى، هل من إمكانية للقول بانتقال هذا الذكاء من التصرف الأخلاقي بأخلاقية، إلى الإدراك الواعي لأخلاقية التصرف الذي قام به من عدمه؟!

أولاً–ذكاء اصطناعي أخلاقي أم يتصرف بأخلاقية!

هذا الطرح البسيط الذي يبين ما يجب أن يكون عليه هذا الذكاء: ذكاءً أخلاقياً أم يتصرف بأخلاقية! يُعيدنا لذات الجدل الفقهي السابق المرتبط بذكاء هذا الذكاء! هل يجب أن يكون الذكاء الاصطناعي ذكياً أم أن عليه أن يتصرف بذكاء؟!(الخطيب، 2021. سبقت الإشارة. الفضلي، 2019. حسين، 2014، ص. ص. 109–132. شبكة، 2012، ص. ص. 19–33. يوسف، 2011، ص. ص. 1–14.) ذلك أن من يعتقد أن الأمرين متطابقان قد جانب الصواب! دون أن يعني ذلك أنهما مختلفان بالكلية، إذ تحكمهما –باعتقادنا– نظرية التكامل لا التطابق.

إذ نعتقد أن الأول ذو مفهوم رياضي مجرد، من خلال تحويل البيانات التي تشكل قاعدة الدماغ الاصطناعي لهذا الذكاء إلى خوارزميات يعمل بموجبها هذا الذكاء بمختلف صنوفه، بما فيها التعلم الذاتي. بينما الثاني ذو مفهوم اجتماعي صرف، يوحي بالضرورة أن ثمة فرضيات ترجيحه، ربما تكون كلها مخطئة أو كلها صائبة، وهنا على الذكاء الاصطناعي أن يختار بينها بذكاء "اصطناعي". بالتالي، إن كان الأول يمكن أن يكون حسابياً أو معيارياً أو حتى قياسياً، كما في حال القول بأنه يجب أن نكون أمام ذكاء ذكي. فإن الثاني المرتبط بالتصرف بذكاء أكثر تعقيداً. إذ أنه يرتبط بالجانب الفلسفي لهذا المفهوم، المرتبط إلى حد كبير بالطبيعة البشرية، وبمعيطات إنسانية واجتماعية وثقافية مختلفة. باعتبارها مرحلة تتجاوز الذكاء إلى القدرة على استخدام هذا الأخير بذكاء. وبالتالي نحن أمام خيارات ترجيحية كلها ذكية، وبالإسقاط على موضوعنا جميعها أخلاقية، وكأننا في إطار اختيار الأصوب بين فرضيات جميعها تعتبر من الناحية الشكلية صحيحة، ولكن عندما تُدمج فيما بينها، يُصبح الأصح هو الصحيح والصحيح هو الخطأ. لننتقل لقاعدة يمكن أن تحكم التصرف الأخلاقي لهذا الذكاء مفادها: أن اختيار الصحيح (الأخلاقي) في معرض الأصح (الأكثر أخلاقية) هو بالضرورة خطأ (غير أخلاقي).

علماً بأن هذا الاختلاف بين أن يكون الذكاء أخلاقياً أو يتصرف بأخلاقية، لا يُثير كبير مشكلة للإنسان، لأننا ندخل هنا ضمن المخزون الأخلاقي لهذا الذكاء، وقدرة هذا الذكاء على التصرف بذكاء ضمن مفهوم الترجيح والاستنساب وقراءة الموقف واتخاذ القرار ... وما إلى ذلك من ظروف استنسابية تُساهم في بناء وتكوين الشخصية الإنسانية ببعديها المعرفي والبيولوجي، بخلاف الحال في الذكاء الاصطناعي الذي وإن اكتملت معرفته، فلا يمكن بحال من الأحوال الحديث عن مفهوم بيولوجي خاص به.

لعل المثال الأكثر توضيحاً لما سبق، يتمثل فيما يعرف بـ: "مشكلة العربة" "Trolley Problem"* (Thomson, 1985. Pp. 1395-1415. Philippa, 1978)، وكيفية التعامل الأخلاقي للذكاء الاصطناعي بين ضررين والاختيار بينهما، والقدرة التفضيلية على اتخاذ القرار المناسب حسابياً ضمن هذا المفهوم في فرضيات يجدر أن تخرج القرارات فيها من الحيز الحسابي النمطي إلى الحيز التفضيلي المبني على المنطق العقلي الإنساني أكثر منه الرياضي الاحتمالي. حيث تُشير هذه المشكلة إلى الصعوبات الوصفية في توضيح السبب أو الأسباب الكامنة وراء التوصل إلى أحكام مختلفة في حالات معينة، تنطوي على أمثلة وسيناريوهات تناقش مُعضلات أخلاقية وفلسفية متعددة (Philippa. Awad, Dsouza, and al. 2018. Pp. 59-64). علماً بأنه لا يوجد مُؤشر واضح لكيفية التعامل مع هذه المشكلة أو المشكلات، بحيث تكون الإجابات متنوعة، بقدر تعدد ارتباطات ذلك بالقرار البشري (Skulmowski, Bunge, Kaspar and Pipa 2014. 8:426. Hagendorff1, 2020 30. Pp. 99-120. P. 99. P. 105).

بالتالي إن مشكلة العربة وطبيعة القرارات المتخذة فيها ليست مسألة تنظيرية صرفة بالمطلق، بل هي على العكس من ذلك تماماً، تمثل حالة تطبيقية بالمعنى الحرفي الدقيق. إذ تتناول ردة الفعل الأخلاقية لهذا الذكاء، تحديداً لمعرفة وفق أي تكييف سيتصرف؟ بمعنى كيف سيفكر الذكاء أخلاقياً وليس حسابياً أو رياضياً. ضمن مواقف غير متوقعة تعرف بـ: "حالات الزاوية" "Corner Cases"، (Bolte, Bar, Lipinski, & Fingscheidt, 2019. Pp.438-445) التي لم يتم تدريب الذكاء الاصطناعي على التعامل معها، في ضوء حقيقة أنه لا يمكن تغطية جميع السيناريوهات التي يمكن أن يتعرض لها الذكاء الاصطناعي، أو حتى تصورها من قبل المصممين البشريين (Bostrom, & Yudkowsky, 2014., p. 316, p. 334. Hazel Si Min; Araz 2019. 11 20: 5791. Pei, Cao, Yang, & Jana, 2017. 1712). من ذلك الحوادث المميتة التي نتجت عن بعض تجارب مركبات Tesla ذاتية القيادة جزئيًا، المرتبطة بسوء تفسير البرمجية الخاصة بهذه

*. مشكلة العربة عبارة عن سلسلة من التجارب الفكرية في الأخلاق وعلم النفس، تتضمن معضلات أخلاقية تدور حول التضحية بشخص واحد لإنقاذ عدد أكبر، وغالباً ما تستخدم في قضايا الذكاء الاصطناعي المرتبطة بالمركبات والعربات ذاتية القيادة، إضافةً للروبوتات لاسيما الاجتماعية. تبدأ السلسلة عادةً بسيناريو يكون فيه قطار أو عربة تم فقدان السيطرة عليها في طريقها للتصادم وقتل عدد من الأشخاص (خمسة أشخاص تقليديًا) على المسار، ولكن يمكن للسائق أو المارة التدخل وتحويل مسار العربة لقتل شخص واحد فقط على مسار مختلف. بحيث يكون لديها خياريين فقط: الأول آلا تفعل شيئًا، وفي هذه الحالة ستقتل العربة الأشخاص الخمسة على المسار الرئيسي. الثاني أن تسحب الرافعة، وتقوم بتحويل العربة إلى المسار الجانبي حيث ستقتل شخصًا واحدًا. السؤال ما هو الخيار الأكثر أخلاقية؟ أو ببساطة أكثر: ما هو الشيء الصحيح الذي ينبغي عمله؟ تبعاً لذلك، يتم طرح أشكال أخرى من العربة الهاربة، ومعضلات الحياة والموت المماثلة الطبية والقضائية وما إلى ذلك. كل منها يحتوي على خيار إما عدم القيام بأي شيء، وفي هذه الحالة سيتم قتل العديد من الأشخاص، أو التدخل والتضحية بواحد في البداية شخص "آمن" لإنقاذ الآخرين. وقد أثيرت مسألة صياغة مبدأ عام يمكن أن يفسر الأحكام المختلفة الناشئة في المتغيرات المختلفة لكل فرضية في ورقة فلسفية عام 1967 من قبل Philippa Foot، وأطلق عليها Judith Thomson "مشكلة العربة" في مقال عام 1976.

العربات للظروف البيئية الفريدة التي لم يتم تدريبها عليها قبل وأثناء الاختبار (Bolte, Bar, Lipinski, & Fingscheidt, 2019, Pp.438-445. Pei, Cao, Yang, & Jana, 2017.).

هذه الفرضية، وكونها ستترك الذكاء الاصطناعي –وفق البعض– ضمن فرضية اتخاذ القرار وحده (Skulmowski, Bunge, Kaspar and Pipa 2014. Bloom. 2011. Pp. 26-43)، ستطرح مستقبلاً تحدياً جديداً يتعلق بتعريف الجيل الجديد من حلول الذكاء الاصطناعي المستقل ذو المنهج الفردي (Autonomous AI solutions with a single approach). بمعنى وفق أي آلية غير متوقعة سيتم اتخاذ هذه القرارات، وضمن أي مجموعة بيانات سيتم التشارك واتخاذ القرار فيها، وأخيراً وهو الأهم وفق أي آلية تفضيلية أو استنسابية سيتم ذلك؟! ما يجعل المسألة ليست أخلاقية ترتبط بالشكل المجرد بهذه المبادئ فقط، بل تتعداها لتتعلق في كيفية بناء هذه المحاكمة الأخلاقية المُفضية لهذه "القرارات الأخلاقية". هنا لا يكمن التحدي في تبني قرار نهائي، وإنما في إنشاء أنظمة قادرة على التواصل وحل المشكلات ذات التفضيلات الأخلاقية المختلفة، وبناء آليات أخلاقية صحيحة من المبدأ أو المبادئ الرئيسة التي تستند إليها. وهو أمر، في ضوء عالمية الذكاء الاصطناعي شديد الخطورة والحساسية، كونه وإن كانت المبادئ الإنسانية متقاربة في الكليات، إلا أنها مختلفة لحد التنافر في بعض الحالات في الجزئيات (الخطيب، ص. 47–90. بسيوني، الدقاق ووزير، 1988. فائق، 2007. الفيلاني، 2007)، وهنا الإشكالية*.

فإن كان الأمان أولاً، هو الشعار الذي تتفق عليه مختلف الشركات العاملة في مجال العربات التقليدية كما الذكية، إلا أن السؤال هو ليس في موضوع الشعار، وإنما في ترجمته بشكل فعلي، حينما تضيق المسافات ويُختصر الزمن في جزء من أجزاء الثانية. ما يطرح قضية الاستقلالية الفردية في اتخاذ القرار في ضوء الغطاء الجماعي لعملية اتخاذ القرار. فالأمر ليس في أننا يجب أن نتجنب الخطر، الأمر المتفق عليه. لكن السؤال هو كيف سنتجنب الخطر؟! وما هي التضحيات المقابلة في ذلك. هنا تدخل القرارات الفردية وتبرز خطورتها. فعلى سبيل المثال، يمكن القول إن على الجميع أن يتلافى الإضرار بالغير. لكن حين تدق ساعة الصفر، كل شخص لديه طريقته وأسلوبه الخاص، الذي يعتقد أنه الأفضل والأصح في الوصول لهذا التلافي والوفاء بهذا الالتزام. ليعمل ضمن ما يعتبره خياراته الصحيحة (Emmanuel; Mekaeel; Danielle; Stefan; Joel; 2012. Pp. 491-498. Simpson, and Hale, 1969. Pp. 495-498).

*. لعل المثال الأكثر قرباً لذلك، واقع الشرعة الدولية لحقوق الإنسان التي تضم ثلاثة مواثيق قانونية دولية هي على التوالي: الإعلان العالمي لحقوق الإنسان لعام 1948، والعهد الدولي المتعلق بالحقوق المدنية والسياسية لعام 1966، ونظيره المتعلق بالحقوق الاقتصادية والاجتماعية والثقافية لذات العام. وأكثر ما تبرز في الإعلان العالمي لحقوق الإنسان، والتناقضات البينية بين العديد من دول العالم حول شرح وتفسير العديد من المبادئ الواردة فيها. فما من قضية عالمية جدلية تثير تخوف الدول ونقاشاتها المحتدمة بقدر قضية احترام حقوق الإنسان في الشرعة الدولية الحالية لحقوق الإنسان، وهل هي معرض توافق عالمي أم تغليب منطق غربي على منطق شرقي، أو تكتل عالمي على آخر! يحاول فرض هيمنته الثقافية ومن خلفها الاقتصادية والسياسية على الآخر. ما ترجم فعلياً من قبل الآخر كردة فعل، من خلال العديد من الإعلانات الجانبية الإقليمية والوطنية، بهدف تحصين نفسه من هذه الشرعة بإعلانات خاصة به لهذه الحقوق.

هنا الإشكالية، التي تتعدى الجانب الأخلاقي في النشاط الإنساني، إلى الجانب الأخلاقي "المصطنع" في الذكاء الاصطناعي. لننتقل من طروحات تتعلق بالعاملين في الحقل التقني إلى أخرى تتعداهم، إلى العاملين في باقي الحقول الاجتماعية والاقتصادية الأخرى.

ففي استطلاع عالمي شهير أجراه معهد ماساتشوستس للتكنولوجيا عبر الإنترنت لمعرفة كيف يجب أن تتفاعل العربة ذاتية القيادة في المواقف الحرجة، لاسيما قضيتي الموت والحياة، أطلق عليه: "الآلة الأخلاقية" "Machine Morale"، توصلت الدراسة إلى استنتاج بسيط مفاده: "ردة فعل هذه العربات تعتمد بالدرجة على بلد المنشأ والمعلومات المتوافرة لديها، خصوصاً عن المشاة والركاب والشبكة الطرقية". وكأننا إن تحدثنا بمفهوم الشخص الطبيعي، إنما نتحدث عن مكان الولادة، وحجم المعارف والخبرات التي يمتلكها، والأخلاقيات التي يختزنها. لذلك إن افترضنا أن عربة ذكية يتوجب عليها أن تتخذ قرارها في اختيار الموت أو الحياة بين ثلاث ضحايا: طفل صغير، ومسن طاعن في السن، ومجرم من أصحاب السوابق. أي من هؤلاء الضحايا الثلاث سوف تستبعد وأيهم سوف تبقي، وبالتالي وفق أي من الخيارات سوف تبني قرارها وأيها سوف تتخذ. هل تتجه لإنهاء حياة المجرم، باعتباره الأكثر استحقاقاً للموت وعقاباً له على جرائمه السابقة! أم سيكون حكمها قائماً على أنه يكفي لهذا الرجل الطاعن في السن من العمر ما عاشه، فهو المستحق للموت! أم تضحي بالطفل باعتباره يمكن أن يكون الأقل تكلفة ...! ... من يحدد؟! (Białek; De Neys. 2016. Pp. 631-639. Bloom. 2011, Pp. 26-43. Unger, 1996. "Famine, Affluence, and Morality". Singer, 2016).

وبعيداً عن ضيق الوقت المتاح لاتخاذ أي من هذه القرارات، إذ نتحدث عن أجزاء الثانية إن لم نقل أقل. الأمر كله لا يتعلق بما نعرف نحن، بل بما تعرفه هذه العربة الذكية. هذه المعرفة التي تتناول المنشأ، المعلومات المتاحة لها، والتي ستختلف بحسب مكان الصنع (الولادة). مع العلم أن هذه الاختلافات، بِقدر ما هي ثقافية بِقدر ما سيترتب عليها نتائج أخلاقية، لها منعكسات قانونية (Sütfeld, Leon; Gast; König; Pipa, 2017. 11: 122).

فعلى سبيل المثال سيكون للعربة الصينية القدرة على الولوج للقضايا ذات الارتباط بالنوع الاجتماعي، والتصنيف الاجتماعي التي تعتبر محظورة في الجانب الأمريكي والأوربي. وبالتالي قاعدة البيانات المتاحة للعربات الصينية المنشأ حتماً ستكون مختلفة عن العربات الأمريكية المنشأ، وكذلك الأوربية، وقواعد البيانات بينهم مختلفة، وبالتالي الأخلاقيات التي تستند إليها كل من هذه العربات، لمعالجة مثل هذه القضايا ستكون حتماً مختلفة، لننتقل من ذكاء اصطناعي وطني لذكاء اصطناعي عابرٍ للحدود. وعليه السؤال لو أن عربة صينية تم شحنها واستخدامها في الولايات المتحدة الأمريكية، وفق أي معايير ستعمل هذه العربة!؟ وفق المعايير الصينية أم الأمريكية! ولنفترض أن ثمة حادث اصطدام سيقع بين عربتين إحداهما صينية والأخرى أمريكية، ما هو البرتوكول الأخلاقي الذي سيتم استخدامه؟! تلك قضية أخلاقية شائكةٌ وشديدة التعقيد.

قضية أخلاقية أخرى يمكن أن تضاف لما سبق، تطرح كذلك في مشاكل العربة، إلا أنها تتعلق بدرجة الذكاء الاصطناعي التي تحوزه هذه العربات. فالتقنيات المرتبطة بدرجة الذكاء قد تكون مختلفة بين العربات ومتفاوتة فيما بينها ضمن ذات المنشأ، بمعنى أنه كلما ارتفعت شريحة العربة زاد ذكائها. فإن كان الجميع يمكنهم قيادة المركبة، فإنهم ليسوا بذات الخبرات المعرفية والمهارات المكتسبة. ذلك أن صنع القرار إنما يعتمد بشكل أساسي على الاعتبارات التقنية. هذا الأمر يطرح إشكالية مدى تداخل الذكاء الإنساني مع الذكاء الاصطناعي في اتخاذ القرار، وكأننا في ضوء الشراكة بين القائد الآلي للعربة والقائد الإنساني فيها وحجم التداخل بينهما، وما هو الحد الأدنى الذي يجب تلبيته في مقتضيات هذا الذكاء. بين مفهوم العربة الذكية بالكامل أو نصف الذكية وما إلى ذلك.

فإن كانت اليوم، شروط الرفاهية التي تحتوي عليها العربات التقليدية لا تثير مشكلة أخلاقية كونها رفاهية ترتبط بالاستخدام البشري والرفاهية التي نسعى إليها في قيادة العربات، فإن الأمر في قضايا الذكاء الاصطناعي ودرجة الذكاء التي يجب أن يتمتع بها، تختلف عن ذلك بشكل كبير، فهي ضرورة وليست رفاهية، وإن اعتبرناها رفاهية فهي رفاهية ترتبط بجودة اتخاذ القرار المرتبط بحماية أرواح كل من يتعامل مع هذه العربات، سواء أكانوا ممن يتواجدون داخل هذه العربة أو خارجها من المحيطين بها. لذلك، نرى أن من الواجب أن تأخذ الأطر التنظيمية هذه القضايا في الاعتبار عند تحديد أين تبدأ مسؤولية صنع القرار الأخلاقي الاصطناعي وأين تنتهي؟

في جميع الأحوال، وإن اعتبرنا –تجاوزاً– أن الذكاء الاصطناعي قد استطاع أن يتجاوز مختلف الفجوات السابقة، فالسؤال الذي يبقى: هل الذكاء الاصطناعي يُدرك عن وعي وتبصر قيامه بذلك. تلك قضية أخلاقية أخرى.

ثانياً–ذكاء يتصرف بإدراك رياضي أم إدراك أخلاقي

الفرضية المثارة هنا تتجاوز التصرف بأخلاق إلى الإدراك الأخلاقي! بمعنى الشعور الواعي باعتبار أن تصرفاً ما هو تصرفٌ أخلاقي. فالسؤال هنا، هل الذكاء الاصطناعي يدرك المفهوم الإنساني للشعور الإنساني لما هو أخلاقي وما هو غير أخلاقي، ولما هو الخير والشر؟! هل يمكننا أن نتخيل يومًا ما تكليف آلة ذكية بمهمة الحكم على الأفعال البشرية، دون أن تكون على دراية بما هو الإنسان وما ينطوي عليه فعل هذا الحكم من نتائج وأثار؟ فعلى سبيل المثال، نعلم جميعنا –نحن والآلة "إن جاز اعتبار ذلك"– أنه لا يجب الكذب، لكن الإحساس بأننا كذبنا تلك مسألة إنسانية صرفة. ما يجعل ميزة معرفة وإدراك ما نقوم به ميزة إنسانية مُفتقدة بالمطلق لدى الآلة (L'interview de Luc de Brabandere, Available online. Sadin, 2018).

عليه الأمر الأهم بعد أن يتصرف الذكاء الاصطناعي بأخلاقية، أن يدرك أنه يتصرف تصرفاً أخلاقياً بأخلاقية، الأمر الذي يتجاوز الأخلاق إلى الوعي الأخلاقي وصولاً للإدراك بالتصرف بأخلاقية وتصويب الأخلاق وتشذيبها وفق مراجعة الضمير. كل ذلك له ارتباط واحد بالإدراك لا بالذكاء. الأمر الذي لا يزال مفتقداً لديها. فهذه الخوارزميات، حتى عندما تؤدي أداءً جيدًا، لا تدرك ما تفعله. ولعل ذلك يبرر عدم وجود أي إشارة لأي من المسائل المرتبطة بوعي الذكاء الاصطناعي والمشاكل الأخلاقية المرتبطة بقضايا الإدراك الاصطناعي (Lyons, 2018)، في إي من الإرشادات الأخلاقية المتعلقة بالذكاء الاصطناعي.

لذلك ربما يكون الأدق والأصلح لهذا الذكاء، الإقرار بإدراكه ما هو الصواب وما الخطأ ضمن المفهوم الحسابي لا الإنساني. مع تأكيدنا بأن القول بأن هناك إدراكا رياضي أو حسابيا ليس صحيحاً بالمطلق، فالإدراك يرتبط بالوعي الذي يفتقد له الذكاء الاصطناعي. كما أن المفاهيم الحسابية بذاتها ليست ثابتة بالمطلق، لاسيما حينما تندرج تحت نظرية الضرورة والمتغيرات الاجتماعية. ضمن ذلك، نعتقد أنه من المفيد جداً حين التحدث عن أخلاقيات الذكاء الاصطناعي أن نراعي في الجوانب الأخلاقية للذكاء الاصطناعي، ليس طبيعة هذه الأخلاق أو حتى امتثال الذكاء الاصطناعي لهذه الأخلاق، وإنما أيضاً التأكد من امتثال القائمين على هذا الذكاء لهذه الأخلاق، ما يُعيدنا لأصل المشكلة والمعالجة معاً: "المبرمجين". لننتقل من أخلقة الذكاء الاصطناعي بذاته إلى أخلقة القائمين على هذا الذكاء. وبالتالي أخلقة الذكاء الإنساني الواضع والمسؤول عن أخلاقيات الذكاء الاصطناعي.

ذلك أنه حينما نُؤخلق أخلقة القائمين على هذا الذكاء، فنحن بالضرورة نُؤخلق أخلقة هذا الذكاء. فكما أن الذكاء الاصطناعي هو بالضرورة امتداد للذكاء الإنساني، فإن أخلاقياته هي بالضرورة امتداد للأخلاقيات الإنسانية وأخلقتها. لنعود لما يؤصل له العديد من أهل الفلسفة التقنية والإنسانية(Le Ny. 1993., p. 12. Yampolskiy & Fox, 2012, Pp. 129-145)، بأننا حينما نبحث في أخلاقيات الذكاء الاصطناعي فكأنما نبحث في أخلاقيات الذكاء الإنساني. إذ إنه من المثبت أنه كلما تعمقنا في الذكاء الاصطناعي وحاولنا فهمه، إنما نغوص بشكل غير مباشر في الذكاء الإنساني. فما محاولة فهم الذكاء الاصطناعي إلا محاولة لفهم ذكائنا الإنساني ذاته. وبالإسقاط، تغدو محاولتنا لأخلقة الذكاء الاصطناعي، ما هي إلا محاولة لأخلقة ذكائنا نحن وفهم حيثياته ودقائقه.

هذا الحيثية الترجيحية في الذكاء الاصطناعي، هي ما تجعلنا ضمن هذه الإشكالية بين القول أن يكون الذكاء ذكاءً أخلاقياً أو أن يتصرف بشكل أخلاقي (Hagendorff, 2020. Pp. 99-120). والحقيقة – باعتقادنا– إن كان الذكاء الاصطناعي ذكياً فهو سيكون أخلاقياً، أما إن كان الذكاء يتصرف بشكل ذكي فسيكون له القدرة على التصرف بشكل أخلاقي. وبالتالي نعتقد جازمين أن مفهوم التصرف بذكاء المُفضي للتصرف الأخلاقي، هو مرحلة تتجاوز أن يكون الذكاء الاصطناعي ذكياً أو أخلاقياً، فهي مرحلة تُبنى على الذكاء والأخلاق، ونحن نحتاج لأن نكون أمام ذكاء يتصرف بشكل أخلاقي لا يخلو من الذكاء. بينما لا

نزال حتى حينه، ورغم كل الجهود في إطار محاولة الوصول إلى ذكاء اصطناعي أخلاقي، بعيدين كل البعد عن القول أنه يتصرف بأخلاقية.

رغم ما سبق، تبقى اللحظة الأهم هي في وصولنا لذلك الإحساس من الرضى بأننا قد حققنا الذكاء الاصطناعي الأخلاقي الذي نريد وننشد. ليبرز السؤال الأهم حول كيفية وصولنا لذلك؟!

الفرع الثاني

كيفية تحقق السلوك الأخلاقي الاصطناعي المُنتظر

إن كان ما سبق يمكن اعتباره سؤلاً فلسفياً عاماً، فإن السؤال الأدق، كيف يمكن لنا أن نتأكد من أننا أمام ذكاء اصطناعي أخلاقي أو ذكاء اصطناعي يتصرف بأخلاقية!؟ الإجابة كذلك ليست يسيرة كما يمكن أن نتصورها للوهلة الأولى. إذ لا يوجد مؤشر أو معيار موثوق وشفاف لهذه الأخلاق أو لما هو أخلاقي من غيره.

عليه، ونظراً لغياب ما سبق، ربما تكون كلمة السر في الموضوع هو تعزيز العمل الإنساني التشاركي وفق توجهين رئيسين: الأول يرتبط بتعزيز أخلاقيات مطوري هذه البرمجيات، والثاني يسعى لتوسيع دائرة التشاركية في المفهوم التكويني للذكاء الاصطناعي وأخلاقياته.

أولاً–تعزيز أخلاقيات مطوري هذه البرمجيات

من المعلوم أن المعتمد عليه في وضع أخلاقيات الذكاء الاصطناعي، إنما يقوم على منهجين رئيسين هما: النهج القائم على الأخلاقيات بالوجوب أو الإلزام، وفق قواعد وواجبات وَاجبة المراعاة. والنهج القائم على أخلاقيات الفضيلة الواجب مراعاتها ضمن هذا الذكاء، فيما يمكن اعتباره "الفضائل الأخلاقية التقنية" "Technomoral Virtues" (Vallor, 2016). بحيث يكون لدينا أخلاقيات قائمة على مجموعة ثابتة من المبادئ التي يجب على مطوري التكنولوجيا الالتزام بها، والتي ترتكز بدورها إلى مفهوم الالتزام بالفضيلة والابتعاد عن الرذيلة. ما يشكل بالمحصلة، خطاب عام يتناول المنظور الأخلاقي وفق مفهوم المبادئ، ومنظور خاص وفق مفهوم الفضيلة يخاطب بشكل فردي مباشر سمات الشخصية والتصرفات السلوكية لمطوري التكنولوجيا.

ضمن ذلك يمكن القول إن التقنيين ومهندسي البرمجيات والمطورين هم المرشحون الأساسيون لمثل هذا الخطاب الأخلاقي، أكثر منه التكنولوجيا نفسها (Mittelstadt, 2019. Pp. 501-507. Ananny, 2016. Pp. 93-117. Leonelli, 2016. Pp. 1-12). ما يوجب تعزيز ما يمكن وصفه بالشخصية الأخلاقية الطبيعة لمطوري هذه البرمجيات، عبر مختلف الفضائل الأخلاقية كالصدق، والعدالة والشجاعة، والتعاطف والرعاية، والكياسة، والشهامة وغيرها. كذلك تعزيز الشخصية الأخلاقية "الافتراضية أو المفترضة أو المُختلقة" لهذا

الذكاء، من خلال ما يمكن تكييفه بالفضائل التقنية كالشفافية، والوضوح والعدالة وعدم الانحياز أو التمييز وغيرها (Vallor, 2016).

ذلك أنه من غير الخافي، أنه يمكن أن يكون للأفعال التي تبدو للوهلة الأولى محايدة أخلاقياً، عواقب أو تأثيرات – مقصودة أو غير مقصودة – خاطئة من الناحية الأخلاقية. ما يستوجب أن يمتلك جميع الممارسين في حقل الذكاء الاصطناعي القدرة على الاستشراف والتمييز في المستتبعات الشاملة –القصيرة والطويلة الأجل– للمشكلات الفنية التي يقومون ببنائها أو صيانتها، إضافة إلى استكشاف طرق بديلة لتطوير البرامج أو استخدام البيانات، بما في ذلك خيار الامتناع عن القيام بمهام معينة تعتبر غير أخلاقية. إذ إن اتخاذ موقف سلبي قد يكون في بعض هذه الحالات تصرفاً أخلاقياً أكثر منه غير أخلاقي.

ثانياً–تعزيز دائرة التشاركية في المفهوم التكويني للذكاء الاصطناعي

من المعلوم أنه لا يوجد بالمطلق علم واحد قائم بذاته، فكما أن الإنسان كائن اجتماعي بطبعه، فإن العلوم –على اختلاف صنوفها– تشاركية في أصلها. لذلك، إن أي منظور تطويري للذكاء الاصطناعي ولأخلاقياته وفق علم واحد، سواء أكان معلوماتي، أو تكنولوجي، أو اجتماعي، أو اقتصادي محكومٌ عليه بالفشل. ما يوجب أن يكون ضمن عملية تكاملية لمختلف هذه العلوم وغيرها. فكما تتضافر الجهود في تكوين الذكاء البشري وأخلاقيته، يجب أن تتضافر الجهود البشرية في تطوير هذا الذكاء الاصطناعي وأخلاقيته. لذلك نعتقد أنه لا ينبغي تطوير الذكاء الاصطناعي وأخلاقياته، بعيداً عن التلاقح الكامل مع باقي العلوم الإنسانية، وفي مقدمتها العلوم القانونية والأخلاقية كما الفلسفية، إضافة للعلوم الاقتصادية والاجتماعية (Bennett Moses, 2011, Pp. 763-94. supra note 58. Koulu, 2020. Pp. 9-46. supra note 37. Larsson, 2019. Pp. 573-93. Supra note 16. Veale, 2020. Pp. 1-8. Supra note 37. Yeung & Martin. 2019)، فجوانب التعلم في الذكاء الاصطناعي هي بالضرورة مدخلات بشرية تُمثل القيم الإنسانية والمجتمعية، وهي التي ستصبح مقياس الأداء الجيد أو السيء لهذا الذكاء. ما يجعل من الضرورة بمكان تحفيز البحث العلمي الإنساني والاجتماعي المرتبط بالذكاء الاصطناعي، بالاشتراك مع نظيره المتعلق بالبحث والتطوير التكنولوجي.

عليه، وضمن مُسلمة الفهم التكنولوجي المُستقرة: "مُدخلات خاطئة مُخرجات خاطئة" " Garbage in Garbage out"، بِقدر ما تكون هذه القيم نبيلة، بِقدر ما نضمن أن تكون المخرجات كذلك، والعكس بالعكس. إذ غالباً ما تكون التشوهات في القرارات الخوارزمية نتيجة لمثل هذه المدخلات. ما يجعل التحدي الأكبر في الفرز المعياري الصحيح، بين البيانات الأساسية وما سينتج عنها من بيانات لاحقة مُستولدة أو مُتضخمة، ضمن ما يعرف بـ: "المسوخات الرقمية" "Deformed digital" أو "التسمم الرقمي" "Digital intoxication" (Artificial Intelligence Index Report 2022. كشفت تجارب CLIP أن صور الأشخاص من ذوي البشرة السمراء تم تصنيفها بشكل خاطئ على أنها غير بشرية أكثر من مرتين مقارنة بالأجناس الأخرى). ما

يوجب إجراء نظرة معيارية دقيقة لهذه البيانات ومدى قدرتها على التكييف والتوسع والتضخم "الجيني" الرقمي غير الخبيث، بحيث تكون أتمتة تقنيات التعلم الذاتي وقابليتها للتوسع أكثر انضباطاً، لتصبح البيانات اللاحقة المُستولدة أو التضخمية، أفضل وأكثر انسجاماً وتوازناً مع البيانات الرئيسية ذاتها، أو تتفوق عليها في المعايير الموضوعية ودرجات الشفافية.

فالذكاء الاصطناعي يتعلم ويكتسب المهارات، ليس من الأمثلة الجيدة والمتوازنة، ولكن كذلك من نظيراتها السيئة التي تعكس الجوانب غير الأخلاقية في هذه القيم، كالعنصرية والكراهية، والتميز، والظلم المجتمعي وغيرها (Larsson, 2019. Pp. 573-93. supra note 16)، ضمن ما يعرف بظاهرة: "المآسي الحسابية" "Algorithmic tragedies"(Martha, 2015. Pp. 122-43).

والأمثلة على ذلك عديدة، إذ يمكن الإشارة إلى برامج الشرطة الفيدرالية الأمريكية لتحديد ظواهر: "التكرار الجرمي" "Criminal Repetition" "Pinpointing repeat offenders"*، الذي ثبت وفق التحقيق الذي أجرته صحيفة ProPublica لعام 2016، بأن البيانات التي تقود هذا النظام المستخدم من القضاة لتحديد ما إذا كان من المحتمل أن يرتكب المجرم المدان المزيد من الجرائم كانت متحيزة ضد الأقليات، لا سيما ذات البشرة السمراء. ورغم أن الشركة التي أنشأت الخوارزمية Northpointe، والمعروفة باسم COMPAS، (Julia, Jeff, Surya and Lauren, 2016) عارضت هذا التقرير والنتائج التي توصل إليها، مُدعية أن أبحاثها أكدت أن دقة نتائج البرنامج المرتبطة بالتنبؤ بمكان حدوث الجرائم تعادل ضعف دقة المحللين البشريين. إلا أن ذلك لم يلغِ حقيقة هذا التحييز المستند إلى العرق.

كذلك، يمكن الإشارة إلى برنامج PredPol المطبق من قبل العديد من أقسام الشرطة في الولايات المتحدة الأمريكية، بهدف معرفة كيف يمكن للتحليل العلمي لبيانات الجريمة أن يساعد في تحديد أنماط السلوك الإجرامي. إذ يحدد هذا البرنامج مناطق بعينها في أحد الأحياء والتي من المرجح أن تحدث الجرائم الخطيرة خلال فترة معينة، والتي غالباً ما تكون ضمن الأحياء الفقيرة أو المهمشة التي يقطنها غالبية من الأقليات العرقية.

مختلف هذه الانحرافات دفعت معهد AI Now المختص بالرقابة على الشفافية في أنظمة الذكاء الاصطناعي إلى التشديد في تقريره الصادر في عام 2017 إلى حد التوصية بعدم استخدام أي هيئات عامة مسؤولة عن مسائل مثل العدالة الجنائية والرعاية الصحية والرعاية الاجتماعية والتعليم أنظمة الصندوق الأسود للذكاء الاصطناعي. مؤكداً أنه نادرًا ما يتم إيلاء الكثير من الاهتمام للمسائل القانونية والأخلاقية

*. حيث طور عالما كومبيوتر أمريكيين أداة برمجية للتنبؤ بالجريمة تسمى: CrimeScan، مستندين لنظرية أن الجريمة العنيفة تشبه في بعض النواحي مرضًا معديًا، يميل إلى الانتشار في مجموعات جغرافية. وبأن الجرائم الأقل خطورة يمكن أن تكون نذيرًا لجرائم أكثر خطورة وعنفًا. حيث يؤكد أحدهم أن: "الاعتداءات البسيطة يمكن أن تتحول إلى اعتداءات جسيمة. أو قد يكون لديك نمط متصاعد من العنف بين عصابتين". لتصبح الفكرة هي تتبع الشرر قبل اندلاع الحريق.

عند إنشاء البرنامج (AI_Now_Institute_2017_Report_.pdf). باعتبار أن تكوين أو تطوير أي برمجية حديثة غالباً ما يكون وفق مبادرات خاصة من أفراد أو جماعات بغرض الاستثمار الشخصي، ما يجعل الكثير من هذه البرمجيات تدخل فيما يمكن اعتباره بالمبادرة التجارية الخاضعة للملكية الخاصة ما يؤثر على موضوعيتها وشفافيتها، ليكون هناك القليل من الشفافية وراء كيفية عمل الخوارزميات (AI_Now_Institute_2017_Report_.pdf).

إضافةً لما سبق، فإنه كلما اتسع نطاق هذه البرمجيات وكلما أصبح التعلم الآلي أكثر تعقيدًا، أصبح من الصعب أكثر بالنسبة للمبرمجين الذين أنشأوا هذا النظام أن يشرحوا أو يفهموا الخيارات التي قاموا بها أو حتى التي يتخذها البرنامج نفسه. ما يؤكد أن تضخم الخوارزميات والتحيز العنصري والجنساني، وكذلك الانحرافات الخوارزمية غير المبررة، إنما هي نتيجة لقيم وأحكام مطوريها البشريين الذين يقررون البيانات التي يجب تضمينها أو استبعادها وكيفية تقييم كل مكون، والتي يصعب عليهم في بعض الأحيان تفسيرها أو شرحها، أو حتى إيجاد المبرر المقنع لها (Drouot, p. 35. Moore, 2017. Bourcier, 2018, p. 8. Godefroy, 2018, p. 734. Chassagnard-Pinet, 2017. P. 495). وبأنه يمكن للبيانات المتحيزة أو غير الكاملة أن تبني نماذج إحصائية مَعيبة وتُعزز التحيزات المجتمعية إذا لم يكن هناك تقييم للأثر أو إجراءات تدقيق أو رقابة مطبقة. وبأن هذا الاعتماد الكلي على الخوارزميات، إنما يؤدي بالمحصلة إلى فقدان القدرة البشرية، خاصة إذا بدأت الخوارزميات في إنشاء خوارزميات جديدة مستولدة (Mihalis, 2018).

لذلك –من جديد– نحن بحاجة لتشاركية فاعلة ومتعددة التخصصات ضمن هذا الذكاء وأخلاقياته. ما يتطلب التعاون بين التخصصات العلمية الحاسوبية والرياضية، إضافةً للهندسية التي لديها رؤى عميقة حول كيفية بناء أنظمة الذكاء الاصطناعي وتشغيلها، والعلوم الإنسانية والتخصصات الموجهة نحو العلوم الاجتماعية والاقتصادية وحتى السياسية التي يمكنها التنظير وفهم مختلف هذه التقاطعات مع الثقافات والمعايير والقيم والمواقف، إضافةً للمعاني والنتائج المترتبة على علاقات القوة في العالم بين التكتلات الاقتصادية والسياسية فيما بينها، وفيما بين الدول كذلك، وصولاً إلى التنظيم الأكثر توفيقاً وتوافقاً لمختلف هذه التداخلات البينية.

وإن كانت المناداة تتعزز اليوم بحضور أهل الاقتصاد والقانون إلى جانب أهل التكنولوجيا في تحديد أطر هذا الذكاء ومحدداته، فإن الوقت مُلِحٌ أكثر من أي وقت مضى بضرورة إقحام أهل الأخلاق ضمن هذا المضمار، بحيث يكون علماء الأخلاق قادرين على استيعاب التفاصيل الفنية من خلال إطارهم الفكري. بما في ذلك التفكير الأخلاقي في طرق إنشاء البيانات وتسجيلها وتنظيمها ومعالجتها ونشرها ومشاركتها واستخدامها، وكذلك طرق تصميم الخوارزميات والرموز، أو حتى طرق اختيار مجموعات بيانات التدريب من أجل تحليل كل هذا بالعمق المطلوب (De Bruin & Floridi, 2017. Pp. 21-39. Kitchin, 2017. Pp. 2018. Pp. 1-17.14-29. Kitchin & Dodge, 2011. Gebru, Morgenstern, & others). وبالتالي الدخول

الأخلاقي في مختلف مراحل تكون هذا الذكاء والانتقال من المفهوم العام للأخلاق إلى مفهوم "الأخلاقيات الدقيقة" "Microethics" التي تنظر في نقاط معينة ومحددة ضمن هذا الذكاء. بحيث نلمس تغييراً جوهرياً في مستوى التجريد والتفكير الأخلاقي يُؤمن وصول هذه الأخلاقيات إلى مختلف التخصصات التقنية وعمليات البحث والتطوير للذكاء الاصطناعي (Morley, Floridi, Kinsey., & others 2019. Pp. 1-21) ، بحيث يكون لدينا: أخلاقيات التكنولوجيا "Technology Ethics"، وأخلاقيات الآلة "Machine Ethics"، وأخلاقيات الكمبيوتر "Computer Ethics"، وأخلاقيات المعلومات "Information Ethics"، إضافة لأخلاقيات البيانات "Data Ethics"، وغيرها من الأشكال التي تتناول بشكل دقيق ومحدد مختلف جوانب هذا الذكاء منذ تكوينه إلى اندثاره. مُكونةً بياناتٍ دقيقةٍ لكل حقل من الحقول السابقة توزع على العاملين في كل حقل، ومن ثم تقاس فيما بينهم وصولاً إلى ضبط أي عملية تحيز أو انحراف غير مبرر. لتكون النتائج المستخلصة من هذه المجموعات من المبرمجين وقياسها ومقارنتها فيما بينهم ومع مجموعات مماثلة، وأخرى غير عاملة ضمن هذه الإرشادات، مؤشراً فاعلاً في تحديد مثل هذه الانحرافات أو الخروقات (Gebru, 2018. Pp. 1-17. Buolamwini, Gebru. 2018. Pp. 1-15.Morgenstern, & others). بخلاف ذلك، وطالما اِمتنع أو مُنع الأخلاقيون عن القيام بذلك، فسيظلون متواجدين في محيط هذا الذكاء لا في الذكاء ذاته، مرئيين بين عامة الناس، لا بين أهل الاختصاص العاملين في المجالات المهنية المرتبطة بهذا الذكاء.

ضمن ذلك تبرز ضرورة وضع مؤشر معياري أو مؤشرات معيارية موثوقة وقابلة للقياس قادرة على تحديد هذه الأخلاقيات، ليبرز السؤال الأدق: هل المؤشر الاجتماعي المستند للضمير يصلح أن يكون مؤشراً قياسياً لأخلاقيات الذكاء الاصطناعي، أم أننا نحتاج إلى مؤشر أكثر موثوقية وصدقية. وهنا، هل يمكن للقانون أن يلعب هذا الدور؟!

المطلب الثاني

مقياس البعد الأخلاقي للسلوك الاصطناعي المُنتظر

السؤال هنا ما هو المقياس الذي يمكن الركون إليه للتأكد من تحقق مفهوم أخلقة هذا الذكاء؟ هل يمكن الركون إلى ما يمكن وصفه بالمقياس الاجتماعي لهذه الأخلاقيات المستند لمفهوم الضمير؟ أم أن الأمر يتجاوز ذلك للمفهوم القانوني؟ أم أن كليهما مُوضع عجز وقصور، ما يقتضي نوعاً من التشاركية فيما بينهما.

الفرع الأول

المعيار الأحادي والصورة المُنتقصة

من المعلوم يقيناً أن الأخلاق مفهوم اجتماعي مختلف عن القانون، يقاس بمقياس اجتماعي صرف هو ضمير الجماعة، فما يوافق ضمير الجماعية من خير هو أخلاقي، وما يخالف ذلك هو شر غير أخلاقي.

وهنا السؤال! هل يمكن تطبيق ذات النظرية في الحديث عن أخلاقيات الذكاء الاصطناعي؟ وإن كان الأمر خلاف ذلك. فهل يمكن للقانون أن يقوم بهذا الدور؟!

أولاً–الضمير الاجتماعي "الاصطناعي" ... المقياس المُفتقد

كما سبق بيانه، بالمفهوم الاجتماعي وبالرجوع إلى مفهوم الأخلاق أو الأخلاقي، ربما التوضيح الأقرب لهما هو المرتبط بالضمير، بمعنى ما يوافق الضمير الإنساني هو أخلاقي وما يخالف هذا الضمير هو غير أخلاقي. غير أن هذا التكييف لا يمكن بحال من الأحوال تبنيه في مفهوم الذكاء الاصطناعي، ليس فقط بسبب طبيعته الضبابية أو حتى عدم موضوعية فكرة الضمير بذاتها لما يتسم به هذا المفهوم من الصفة الشخصية لا الموضوعية، ما يتناقض مع التكييف الأخلاقي لمفهوم عالم عابر للحدود يتناول الحياة البشرية والإنسانية برمتها، بل لانعدام وجوده بالأصل ضمن –فيما لو جاز لنا التعبير– "مجتمع الذكاء الاصطناعي".

ذلك أن قضية الضمير ترتبط بالمجتمع الذي يُعَبر عنه هذا الضمير، والذي تبلور فيه مفهوم هذا الضمير. وهو أمر غير ممكن القول بتحققه في الذكاء الاصطناعي. فأخلاقيات الذكاء الاصطناعي ليست من صنع هذا الذكاء. فنحن بعيدين كل البعد عن مجتمع قوامه الذكاء الاصطناعي الذي يحدد ضميره أخلاقيات هذا الذكاء لديه. إذ لا يمكن القول بأي حال من الأحوال أننا أمام جماعة رقمية أو آلية لها ضمير خاص بها، يضع لها مجموعة من الأخلاقيات التي ترتكز إليها. ذلك أن من يحدد هذا الذكاء هو ضمير الجماعة التي صنعت هذا الذكاء، وفق أخلاقيات الجماعة الإنسانية التي تتبع لها، وبالتالي حين الحديث عن مقياس لأخلاقيات الذكاء الاصطناعي نحن نتحدث عن مقياس لأخلاقيات الجماعات البشرية بمختلف مسمياتها التي تضع هذا الذكاء، وليس مقياس "اجتماعي مصطنع" لهذا المجتمع. فالمفاهيم الإنسانية التي تُخَلِقُ الأخلاق وتُؤطرها، ضمن حدود الوعي والتعاطف وفهم اللغة، وتقدير الموقف وغيرها ... والمرتبطة بالخصائص البشرية الإنسانية ذات البعد الروحي أو الشعوري أكثر منه الجسدي أو الفيزيائي، لا تزال غير متحققة أو قابلة للقياس فعلياً ضمن الذكاء الاصطناعي، وإن تحققت فهي مجرد تمثيل رياضي أو خوارزمي، وهنا لا نقصد مجرد القول بما هو جيد وما سيئ، بقدر ما الإحساس أو الشعور بأي منهما، الأمر المفتقد ضمن هذا الذكاء.

هذا الافتقاد المطلق لمفهوم الضمير الاجتماعي "الاصطناعي" ينقلنا لفرضية أكثر تعقيداً مفادها: هل يمكن الركون في تحديد هذه الأخلاقيات إلى ضمير الجماعة "الإنسانية" التي كونت هذا الذكاء؟!

هنا السؤال كذلك يحتاج إلى عمق وتحليل، وإجابته باعتقادنا كذلك النفي! فعدا عن العيوب التي تعتري مفهوم الضمير بشكل عام، من حيث الضبابية وعدم الموضوعية. فإنه من المعلوم أنه كذلك مفهوم غير ملزم ولا يمكن قياسه، كما أنه يخضع للاختلاف من جماعة إلى أخرى، لاسيما حينما تُدمج فيه مُقتضيات

الضمير الفردي بالضمير الجماعي أو تُخالفها. علماً بأن هذه التباينات ستبرز بشكل أكثر وضوحاً وتعقيداً حينما ننتقل من مفهوم الضمير المجتمعي لجماعة أو مجموعة محددة من الأفراد إلى مفهوم الضمير المجتمعي للإنسانية جمعاء ضمن مفهوم عالمي عابر للحدود كما هو الحال في أخلاقيات الذكاء الاصطناعي.

وهنا تغدو الأمور أكثر تعقيداً وصعوبة! ليغدو الحديث عن مقياس اجتماعي لأخلاقيات الذكاء الاصطناعي وفق هذا المفهوم أمرا خارجا عن المتصور ولا يمكن الركون إليه، ليبرز الحديث عن دور ممكن لعامل أكثر فاعيلة في هذا المجال وفي تحديد هذه الأخلاقيات هو "القانون"، باعتباره القادر على أن يكون مقياسا لهذه الأخلاقيات ومقياس التصرف بأخلاقية. بمعنى أن ما يتوافق مع القانون هو أخلاقي وما يتعارض معه غير أخلاقي. رغم ذلك، يبقى السؤال: هل هذا التوجه هو الخيار الأمثل؟!

ثانياً–الالتزام القانوني الصرف ... المقياس المُنتقد

أمام ضَعف الرأي المؤيد للأخذ بالمعيار الاجتماعي نظراً للاعتراضات المُحقة السابقة، برز التوجه القائم إلى الأخذ بالقانون كمعيار فاعل لقياس أخلاقيات هذا الذكاء. توجهٌ رغم موضوعيته والترحيب الذي يمكن أن يحوزه شكلياً، إلا أنه لا يخلو من الانتقاد موضوعياً، لأسباب عديدة.

فالأخذ بهذا المعيار سيجعلنا نتعامل مع مفهومين مختلفين: "الأخلاق والقانون" وكأنهما مفهومان متماثلان، علماً بأن أولى الأدبيات القانونية تشدد على الاختلاف التأصيلي بين المفهومين. كما أنه يجعلنا أمام تلاعب في العبارات والمفاهيم، كون أخلقة الذكاء الاصطناعي في هذه الفرضية ما هي إلا تعبير عن تقنين هذا الذكاء. وهو الأمر الذي لا ينسجم مع الطرح الذي نريده من أخلقة هذا الذكاء. كذلك، فإن النظرة إلى الأخلاق وفق معيار حسابي لها مخاطرها التي يجب أن تؤخذ بالحسبان، بمعنى إن كان مفهوم الخير ومفهوم الشر هو مفهوم الأغلبية أو ما ستراه الأغلبية كذلك (Michael & Susan Leigh. 2007, Pp. 15-26. Ștefan, 2020, Pp.136-148. Dignum, 2019) ، فإن هذا الأمر لا شك سيفضي إلى الإضرار بالأقلية على حساب الأغلبية وفق الحساب الأخلاقي. ذلك أن ما نعتبره أخلاقيا ليس الأخلاقي بذاته، وإنما ما ستراه الأغلبية أخلاقيا! هذا الأمر سيجعل حقوق وحريات الأقليات في معرض الذكاء الاصطناعي مُنتهكة بشكل أخلاقي، ضمن مفهوم الجبر والإلزام. الأمر المُنتقد.

لذلك، ربما يكون مفيداً وأكثر إنصافاً، الأخذ بالمفهوم الأخلاقي وفق المفهوم المعياري، وفي ضوء التصنيف الأخلاقي للأفعال بين ما هو مسموح وممنوع واختياري وإلزامي (Ștefan, Pp. 136-148). رغم أننا نعتقد أن هذا التداخل لا يزال قائماً بين المفهومين في المفهوم الأخلاقي الإنساني، ويزداد تعقيداً فيما يمكن وصفه بالمفهوم الأخلاقي الاصطناعي. ذلك أن الوصول إلى ما هو مسموح أو ممنوع، إنما يستند بالضرورة إلى معيار إنساني بامتياز، وفق ما يعتبره الضمير المجتمعي كذلك. وإن كان للفكر الإنساني

القدرة على التمييز الواعي بين مختلف هذه التصورات، باعتباره قادرا على فهم مكنونات البعد الإنساني لهذه الأخلاق المرتبط بالضمير، بشكل يجعل هامش الخروج عن هذه المعادلة الحتمية ممكناً في الذكاء الإنساني، لأنه يمتلك الحس الإنساني المُفضي لمعرفة وتحديد ما هو خير وما هو شر وفق تقيمه الشخصي بعيداً عن رأي الجماعة أو أي أمر أخر. فإن هذه الفرضية غير ممكنة بالمطلق في الذكاء الاصطناعي، كونه يستند بالضرورة إلى المعيار الحسابي المفضي للضمير الاصطناعي الجماعي المُفتقد، مع الغياب الكلي لكل ما يمكن وصفه بالضمير الاصطناعي الفردي الذي يسمح له بتقييم حساب ضمير الجماعة ضمن ضمير الفرد والخروج بمخرجات فردية ترضي الضمير الفردي على حساب الضمير الجماعي. كل ذلك مُفتقد بالمطلق في الذكاء الاصطناعي.

الفرع الثاني

التآزر الأخلاقي والقانوني والصورة الكاملة

في ضوء قصور أي من المعياريين السابقين في ضمان أخلقة ناجعة للذكاء الاصطناعي، يكمن الحل –باعتقادنا– في ضمان التآزر بينهما كوسيلة تكفل الوصول إلى الحل الأخلاقي الأمثل لهذا الذكاء ضمن المفهوم التأصيلي أولاً والاستشرافي ثانياً، بكونه الضامن الفعلي المُستدام لهذه الأخلاقيات.

أولاً–التآزر ... الحل التأصيلي الأمثل

كل حديث اليوم عن أخلاقيات الذكاء الاصطناعي هو حديث عن أخلاقيات القائمين على الذكاء الاصطناعي لا الذكاء الاصطناعي بذاته، ولكن وفق مفهوم قانوني صرف لا اجتماعي أو فلسفي بالمطلق. ذلك أن جميع ما يرتبط بالذات الإنسانية والمكونات الاجتماعية هي خاصة بالمكون الإنساني المرتبط بالشخصية المفتقدة بالمطلق لدى الذكاء الاصطناعي وفق المستقر والثابت في هذا الموضوع (Bensoussan, 2013, n° 1134. Bourcier, 2001, Pp. 847-871. Loiseau, 2018, Pp. 1039-1042. | الخطيب، 2017، "سبقت الإشارة. ص. ص. 19–58. المناصير، 2020، ص 44.). ما يجعل فكرة التآزر بين الأخلاق والقانون أولوية في تأطير أخلاقيات هذا الذكاء (Langlois, 2019). فبقدر ما توفر الأخلاق منطقة جذب طوعية تسمح للقائمين على هذا الذكاء بالتكييف بشكل أكثر مرونة مع إحساس أقل بالإلزامية وشعور أكبر بالحرية، بقدر ما يمكن للقانون أن يعطي مختلف هذه التوجهات الأخلاقية طابعاً إلزامياً غير مباشر من خلال المراقبة والمتابعة (Dermange et Flachon, 2002, P. 15). ليكمل كل منهما الآخر. بحيث أن ما يسكت عنه القانون تُكمله الأخلاق، وما تعجز عنه الأخلاق ينهض به القانون. وبذلك سيكون كل من الأخلاق والقانون مشاركين في تنظيم الذكاء الاصطناعي. بحيث تنتقل الحقوق الأساسية من القانون إلى الأخلاق، بينما تنتقل الشفافية والإنصاف واستقلالية الإنسان من الأخلاق إلى القانون.

هذا الأمر –فيما لو طبق– سيسمح بالتجاوز الإيجابي لحدود القانون ضمن حدود الأخلاق، فإن كان مقولة أن ما هو قانوني هو أخلاقي وما هو خلاف ذلك هو غير قانوني استنتاج صحيح من الناحية القانونية، إلا إنه يبقى منقوصاً من الناحية الأخلاقية. بمعنى أن القانون لا يمكنه أن يُحيط بكل شيء، سيما وأن الأصل في سياسة المحاسبة المدنية لا تقوم على النص والتقنين، وبالتالي حينما يعلم الشخص أن تصرفه غير أخلاقي، وإن لم يكن غير قانوني كونه ليس هناك نص يقضي بالمعاقبة عليه، لا لأن القانون يسمح به، بل لأن القانون لم يتصوره! هنا نعتقد أن الشخص قد تجاوز المعيار الأخلاقي رغم كونه لم ينتهك المعيار القانوني، ما يجعل المساحة الفاصلة بين القانون والأخلاق موجودة ومتأصلة أيضاً في الجوانب الأخلاقية التي يعلمها ضمير الشخص أو الجماعة ولم يتناولها القانون. كما أن هذا التآزر الأخلاقي والقانوني سيمكن من اعتبار أن الأخلاقيات المُؤطرة قانونياً إنما تمثل الحد الأدنى من الأخلاقيات الذي لا يجوز النزول دونها. بينما فيما لو أرادت بعض الشركات أو القائمون على سياسات الذكاء الاصطناعي وضع أخلاقيات أكثر تشدداً بمعايير أخلاقية أعلى، فإنه لا شيء يمنع ذلك. بذات الوقت الذي يغدو هذا التوجه موضع التزام أخلاقي قانوني خاص بهذه الجهة أو الفئة.

استنتاجاً مما سبق، يمكن القول أن مفهوم أخلاقيات الذكاء الاصطناعي، هو بالضرورة مفهوم حسابي أو معياري يتناول أخلاق الذكاء الاصطناعي وفق مؤشرات معيارية محددة يجب على القائمين على الذكاء الاصطناعي الالتزام بها تحت طائلة المؤيد القانوني، ما يجعلنا أمام دمج قسري محمود تفرضه طبيعة الموضوع المعالج بين الأخلاق والقانون، ليغدو القانون هو المعبر عن ذاته وعن الأخلاق معاً. ما يجعل الاستنتاج بأن أخلاقيات الذكاء الاصطناعي هي قوانين الذكاء الاصطناعي، استنتاجا لا يخلو من الصحة والمصداقية.

وهنا من الواجب علينا نحن رجال القانون أن نعيد التأكيد بأن الهدف من أخلقة هذا الذكاء، هو جعله أكثر تفهماً لاحتياجات الناس وأكثر قرباً منهم من خلال تعزيز مفهومي الشفافية والمساءلة فيه، إضافةً إلى الوقوف على نقاط العتمة الرقمية فيه، وتصويب انحرافاته الخبيثة والإشادة بخصائصه الحميدة (Pariser, 2011. The filter bubble: 2018. About us. 2019. Howard, Kollanyi. 2016. Bots, #StrongerIn, and #Brexit. Pp. 1-6. Vosoughi, Roy, & Aral. 2018. Pp. 1146-1151. Brady, Wills, Jost, Tucker, & van Bavel, 2017. Pp. 7313-7318.)*. ذلك أنه إن كانت الخوارزميات وما يعرف بالتقنيات المزعجة "Disruptive Technologies" (Taeihagh, Ramesh, & Howlett, 2021) قد اتحدتا لتبيان بعض جوانب القصور في تكييف القانون مع المستجدات المتسارعة حوله، محاولةً بذلك تغير العديد من مفاهيمنا القانونية.

*. كما في قدرته على تحديد الاستخدامات غير الأخلاقية لبعض البرمجيات كما في حال تأثيرات غرفة الصدى، أو استخدام الروبوتات الدعائية، أو نشر الأخبار المزيفة، أو محاربة التطرف، وخطابات الكراهية المفضية للانقسام المجتمعي. أو تصميم الشبكة الذكية للحد من آثار أزمة المناخ، وتطوير البنية التحتية منخفضة الانبعاثات، ونمذجة تنبؤات تغير المناخ. ضمن جميع هذه الفرضيات لعب الذكاء الاصطناعي دورًا فاعلاً في الحد من مختلف هذه الاستخدامات غير الأخلاقية.

فإنه علينا نحن أهل القانون، أن نعمل مجدداً لإيصال صوت القانون وغايته والمهمة التي ينهض بها للجميع بمن فيهم التقنيين والمبرمجون. كما أن علينا ألا نشعر بالضعف أو الخجل، أو ننقاد أو نبقى مرتهنين لما يقوله التقنيون، أو يأتينا منهم، وأن نبتعد عن السرد المشتت غير المجدي بأن الابتكار يقود كل شيء وأن الرمز (Code) هو القانون، وأن كل شيء متأخر، أو يتبعه (Lessig, 1999. Lessig, 1999. Floridi. 2018). بل إن علينا أن نؤمن بأن الاستماع لأصوات منتقدي المنهجية القانونية التقليدية بمن فيهم من هم من أهل القانون ومن المبرمجين والتقنيين، إنما هو وسيلة جيدة لإعادة فهم القانون من جهة وللرد على المشككين في قدرة القانون على التكييف من جهة أخرى. وبأن الأمر برمته يجب أن يكون في خدمة القانون لا العكس(Lessig, 2000)*.

فالسؤال هنا: إن كنا لن نسمح للذكاء الاصطناعي أن يحدد قيمنا فهل سنسمح للمبرمجين تحديد هذه القيم (Dignum, 2019. | الخطيب، 2020، المقال الرابع. الخطيب، 2018، ص. ص. 251-324.). ما يدفعنا للتسليم بحقيقة واحدة مفادها: "أن من يقود العربة هو القانون لا سواه" (Oster, 2021, Pp. 101-117. | Lessig, Godefroy., p. 734. الخطيب، 2020، المقال الثالث. الخطيب، سبقت الإشارة.).

كذلك، إن كان الذكاء الاصطناعي لا سيما المستند للتعلم العميق، لا محالة سيعد النظر في العديد من المصطلحات القانونية التقليدية مثل الاستقلالية والخصوصية (Dignum, 2019,. الخطيب، سبقت الإشارة. الخطيب، سبقت الإشارة، ص. ص. 251-324.). فإننا بحاجة لإعادة التأكيد بأننا أمام مصطلحات قيمية لا تقنية. وأن هذه المصطلحات، وإن أدخلت كرموز ضمن برمجية الذكاء الاصطناعي، إلا أنه لا يمكن التعامل معها على أنها مجرد مسائل هندسية أو رقمية (Mihalis, 2019). فكما أننا لا نثق في أي قاضٍ لبناء شبكة عصبية عميقة، يجب أن نتوقف عن افتراض أن شهادة الهندسة كافية لاتخاذ قرارات معقدة في مجالات القانون والأخلاق.

فالقانون ليس مجرد عملية حسابية أو رياضية، بل عمل إنساني بامتياز، كما أنه ليس مجرد نبوءة رقمية أو اصطناعية (Foss-Solbrekk, 2021, Pp. 247-258. Gaudemet, 2018, p. 651. Meneceur, 2018, p. 316. | الخطيب، 2019، ص. ص. 11-48. الخطيب، 2019، المقال الثالث. مرقس، 2014. إبراهيم 2008.)، بل هو نبوءة إنسانية (Oliver, 1997: Pp. 991-1009. Glastra van Loon, 1958: Pp. 1-9.

*. « The code of cyberspace regulates». « Our choice is not between "regulation" and "no regulation". The code regulates. It implements values, or not. It enables freedoms, or disables them. It protects privacy, or promotes monitoring. People choose how the code does these things. ... The only choice is whether we collectively will have a role in their choice - and thus in determining how these values regulate - or whether collectively we will allow the coders to select our values for us ». « The law of cyberspace will be how cyberspace codes it, but we will have lost our role in setting that law ». Chassagnard-Pinet, , p. 495. « L'emploi des algorithmes dans le processus de production du droit pourrait donc être à l'origine d'une normativité juridique d'une nature nouvelle : une normativité immanente, personnalisée, non délibérée ».

| Frison-Roche, art. 4. Favoreu, 2002, p. 513. الخطيب، 2019، ص. ص. 267-311. 2019، ص. ص. 117-180. يوسف، 2012، ص. ص. 1079-1106. فايد، 2012، ص. ص. 961-999.). كما أن القانون يتجاور مفهوم المنطقية إلى التجربة، وتحديداً التجربة المتغيرة وليس الثابتة. ضمن ذلك فإن مفهوم الثوابت لا يصلح مطلقاً للقانون. كذلك إن القانون ليس مجرد عملية تجميع معلومات قانونية، بل جوهره يدور حول كيفية التعامل مع هذه المعلومات بالتحليل والاستنتاج. فالمعرفة القانونية لا يمكن اختزالها إلى معلومات عن الحلول المطبقة أو قابلة للتطبيق، ولكن يعتمد بشكل أساسي على الطريقة التي يتم بها طرح المشكلة القانونية وتسليط الضوء عليها (Rouvière. 2017. P. 527, spéc. 528, 2e col). ما يعيد طرح التمييز بين النظرية الرياضية للمعلومات التي تؤسس كلاً من علوم الكمبيوتر والذكاء الاصطناعي عن النظرية الإنسانية التي تأسست عليها النظريات القانونية وفق ما هو مؤسس عليه، وهو ما يتماشى مع المغزى الإنساني لفكرة القانون (Sommer, Oster, , Pp. 101-117. Lessig, . Godefroy, , p. 734. Rouvière, 2020, p. 990. 2021.Blin-Franchomme, 2021 p. 100. Castets-Renard, 2021 p. 297. Larrière, 2017, Pp. 38-40. Quenillet, 1994, p. 11 Godefroy, , p. 734. Drouot, 2020. p. 35. | Chatila, 2016, p. 284. Hildebrandt, 2017, Pp. 579-608. الخطيب، سبقت الإشارة. الخطيب، سبقت الإشارة. الخطيب، سبقت الإشارة.).

علماً بأن هذه القضايا على اختلاف تنوعها تواجه تحدياً يتعلق لما يمكن وصفه بالفجوة الرقمية بين الأجيال "Intergenerational digital divide" والنظرة الأخلاقية لكل ما سبق، وذلك بحث آخر (Mihalis, 2018).

ثانياً-التآزر ... الحل الاستشرافي الأمثل

لا نضيف شيئاً إن قلنا، إن الذكاء الاصطناعي وتداخلاته في حياتنا اليومية أصبح اليوم كما الهواء الذي نستنشقه، إذ أنه يُحط بنا من كل جانب، سواء أحسسنا به أم لا. لكننا لعلنا نستطيع الإضافة حينما نشير إلى الضرورة الحتمية لتنقية هذا الهواء من جميع الشوائب، بما يجعله عاملاً مساعداً على الحياة، لا ضاراً بها. كونه ليس بنظيف بالكلية، ولما لذلك من ارتباط وثيق بين أخلقة هذا الذكاء وتنقيته، ما يجعل التآزر الفعلي بين القانون والأخلاق مسألة ثقافية واجتماعية مُستدامة ومُستمرة، نعود فيها لمبدأ الحيطة والحذر لتجنب التعامل مع مخاطر هذا المسار التكنولوجي سريع التغير.

ضمن ذلك، نرى أن أخلقة الذكاء الاصطناعي في مرحلة التكوين أشد أهمية منها في مرحلة الاستخدام. ما يجعل القول بأن أخلقة الذكاء الاصطناعي، إنما تكمن في الاستخدامات فقط قول يفتقد للدقة. فالأخلقة تكمن في التكوين ولا نغالي إن قلنا أنها تكمن في التكوين أولاً وأخيراً. ذلك أن تحصين الذكاء الاصطناعي، بالأخلاقيات عند التكوين إنما يساعد كثيراً في الحد من الاستخدامات السيئة بعد التكوين، وربما يمنعها.

فكأنما نحن أمام طفل أُحسن تهذيبه "ترتبيه"، وآخر أسيء تهذيبه "تربيته". لذلك نعتقد جازمين، أن قضية أخلقة الذكاء الاصطناعي، إنما هي أقرب لقضية تهذيب الذكاء الاصطناعي وتنشئته النشء الصحيح.

وإذ نقر بأن قضية التهذيب اليوم ضمن مفهوم الأخلقة قضية ليست بالهينة، لما يكتنف عملية تكوين هذا الذكاء من تعقيد وغموض، ولما يتداخل في هذه العملية من مهذبين "معلمين" متعددين من خلفيات وثقافات متعددة، ناهيك عن صعوبة الحديث عن هذه الأخلاقيات، إن من حيث المضمون الفلسفي التأصيلي لهذه الأخلاق، أو من حيث طبيعة هذه الأخلاق التي لا تخضع لمعيار رياضي وإنما لجملة معايير اجتماعية ترتبط بالمبادئ الإنسانية العالمية المتفق عليها، وكذلك المختلف حولها أو حول مضمونها الدقيق(كما سبق بيانه، المثال الأكثر قرباً لذلك، واقع الشرعة الدولية لحقوق الإنسان). إلا أننا نشدد على أنها المخرج الوحيد لهذه الحوكمة. ما يدفعنا للاستنتاج، بأن أخلقة الذكاء الاصطناعي وصولاً لحوكمته، إنما هي أقرب لمفهوم تربية الذكاء الاصطناعي وصولاً إلى محاسبته. بمعنى جعله ذكاءً أخلاقياً، قابلاً للمحاسبة في حال خروجه عن حدوده الأخلاقية، مع بقائه ضمن مفهوم القاصر لا الراشد. بمعنى أنه حال خروجه عن المحددات الأخلاقية سوف يسأل المسؤول عنه، ضمن مفهوم التبعية. في ضوء هذه المعادلة، يغدو الحديث عن الأخلقة مَدخلاً رحباً للحديث عن مفهوم محددات هذه الأخلقة والمسؤولية المترتبة عليها(Ştefan, , Pp. 136-148)*.

وهنا نؤكد كذلك، أن دورة حياة الذكاء الاصطناعي –إن جاز لنا استخدام هذا التعبير– سيكون لها دور فاعل في تحديد طبيعة أخلقة هذا الذكاء والأدوار التي يمكن أن تنهض بها مختلف الفئات والجهات المؤثرة في ذلك. ذلك أنه ضمن هذه الدورة "الحياتية" للذكاء الاصطناعي، يمكن التحدث عن أربعة مراحل مفصلية: بين الإنشاء والتكوين مروراً بالمراجعة والاختبار، وصولاً للإنتاج والتصنيع، وانتهاءً بالنشر والتشغيل. ففي التكوين، يمكن الحديث عن تصميم نظام الذكاء الاصطناعي، عبر تحديد الهدف منه والنظام الذي سيعمل عليه، ونمط النموذج ومقاييس الأداء والبيانات ذات الصلة أو المدخلات الأخرى المتعلقة به، بما فيه ما يرتبط بجمع البيانات وتنظيفها وفحص الجودة (بما في ذلك التحيز) والتوثيق. بينما تنطوي المراجعة والاختبار، على العلميات المرتبطة بتقييم الأداء الخاص بهذا الذكاء، بما ينطوي عليه من استكشاف لأخطاء البرمجيات وجوانب التحيز ونقاط العتمة، ومقارنة سلوك النموذج السابق بالسلوك الجديد، والأداء عبر العديد من المقاييس، كمقاييس الدقة والمعايرة. في حين أن مرحلة الإنتاج أو التصنيع، تتعلق بالرقابة والمتابعة ذات الصلة بتشغيل الذكاء في حالات واقعية بغرض التحقق من الامتثال للوائح ذات الصلة، والتحقق من التوافق مع البرامج القديمة، وتعيين المسؤوليات لإدارة نظام الذكاء الاصطناعي. وأخيراً، تأتي مرحلة النشر

*. The importance of learning and controlling the problem of ethicality of replies generated by a conversational agent trained with ML was emphasized in the case of the Tay bot. Microsoft delivered this bot on Twitter for entering in dialog with people but was forced to remove it only after 16 hours because its language was abusive and offensive, it became racist and mysoginist in its replies. Available online, retrieved on March 6, 2021.

والتشغيل التي تتضمن طرح الذكاء في السوق وتقييم مخرجاته وتأثيراته بناءً على نوايا المصممين الأولية ومقاييس الأداء بالإضافة إلى الاعتبارات الأخلاقية. ليتم من خلال هذه المراحل الأربعة تحديد المشكلات ومعالجتها ضمن هذا الذكاء، إما من خلال العودة إلى المراحل الأخرى، أو من خلال اتخاذ قرار راديكالي، يتمثل بإلغاء نظام الذكاء الاصطناعي برمته.

وإذ نُشيد بالجهود المبذولة في مجال أخلقة هذا الذكاء، لاسيما بالنسبة للمعايير القابلة للقياس الرقمي، كما هو الحال في حماية الخصوصية والبيانات، ومكافحة التمييز، والسلامة (Baron, Musolesi, 2017. Mittelstadt, Russell, Wachter. 2019. 2018. Pp. 1-17..Gebru, Morgenstern, & others Pp. 1-10. Pp. 1-10)*، أو التفسير. نُؤكد أن ثمة جوانب أخرى مُغفلة، تم تجاهلها إما عن قصد أو غير قصد، كما هو الحال في خطر الذكاء الاصطناعي العام الخبيث، وقضايا الوعي الآلي، والبرمجيات المرتبطة بالحد من التماسك الاجتماعي من خلال أنظمة التصنيف والتصفية بالذكاء الاصطناعي على مواقع الشبكات الاجتماعية، والاعتداء السياسي لأنظمة الذكاء الاصطناعي، ونقص التنوع في مجتمع الذكاء الاصطناعي، والمسائل المرتبطة بروابط أخلاقيات الروبوتات الاجتماعية، والتعامل مع مشاكل العربة، والوزن بين الخوارزميات أو إجراءات اتخاذ القرار البشري، والتكاليف الاجتماعية والبيئية "الخفية" للذكاء الاصطناعي، ومشكلة الشراكات بين القطاعين العام والخاص والبحوث الممولة من الصناعة ... والقائمة تطول (Rahwan, 2018. Pp. 5-14. Leiser, & Murray, 2016. Pp. 670-704. Dhar, 2020. Pp. 423-425. Crawford, Joler. 2018. Kate, and Vladan, 2018. Irani, 2016. Pp. 34-37. Veglis, 2014). مع التأكيد بأن قائمة هذه الإغفالات تختلف المبررات فيها، إذ لا يمكن القول بأن جميع مبرراتها مقنعة. فمنها ما يرتبط بطبيعة هذا الذكاء وتطوره وعدم القدرة على الاستشراف المستقبلي الكلي والواضح لتطوره، كما هو الحال في المداولات حول الذكاء الاصطناعي العام، من خلال الإشارة إلى طبيعتها التخمينية البحتة، في حين أن العديد من الإغفالات الأخرى أقل إقناعاً، ويجب أن تكون سببًا لتحديث أو تحسين الإرشادات الحالية والمقبلة (Graham, Hjorth, & Lehdonvirta, 2017. Pp. 135-162. Silberman, Chen. 2017. Tomlinson, La Plante, Ross, Irani, & Zaldivar. 2018. Pp. 39-41. Irani, 2015. Pp. 720-739. Irani, Pp. 283-314,2016., Pp. 34-37. Veglis, 2014, Pp. 137-148. Fang, 2019. 2020).

*. L. n° 2016-1321 du 7 oct. 2016 pour une République numérique, JO n° 0235 du 8 oct. Règl. UE 2016/679 du Parlement européen et du Conseil du 27 avr. 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données (RGPD), JOUE L 119/1, 4 mai. L'Union européenne est fondée sur l'engagement constitutionnel de protéger les droits fondamentaux et indivisibles des êtres humains, de veiller au respect de l'état de droit, d'encourager la liberté démocratique et de promouvoir le bien commun. Ces droits sont reflétés aux art. 2 et 3 du Traité sur l'Union européenne, ainsi que dans la Charte des droits fondamentaux de l'Union européenne. D'autres instruments juridiques reflètent et précisent ces engagements, comme la Charte sociale européenne du Conseil de l'Europe.

وعليه نرى أن هذه العيوب الرقمية بمختلف مسمياتها وطبيعتها، لا سيما المتولدة من افتقاد التنسيق والتكامل بين العلوم الإنسانية والعلوم التقنية، إنما مردها لديناميكيات القوة بين الجهات الفاعلة الموجودة بشكل مستقل عن الأدوات التقنية المستخدمة. ولكونه لا مكان للفراغ في مختلف هذه العلوم، فإنه إن لم تستطع العلوم الإنسانية ملء الفراغ المتوجب عليها في هذه العلوم التقنية كما هو واجب ومطلوب، ستقوم العلوم التقنية بملء هذا الفراغ وفقاً لفلسفتها، بغض النظر عن النتائج التي يمكن أن تترتب على ذلك. لذلك يجب فهم هذا الذكاء ضمن منظور ينظر بالكلية لمختلف العلوم التي تساهم في صناعته وتكوينه، فكما أنه لا يمكن فهم أي قطعة أثرية بشكل صحيح دون الرجوع إلى سياقها التاريخي والإنساني، لا يمكن فهم أنظمة البرامج دون ربطها بمختلف العلوم التي يمكن أن تتأثر وتؤثر بها (Kroll. 2018, p. 11).

ضمن ذلك نعتقد أن الاستجابة لهذا الذكاء تتطلب المزيد من الدراسات البينية بين مختلف العلوم الإنسانية بشقيها الاجتماعي والعلمي، والاستماع لهموم الجميع ومراعاة مصالح الجميع، خاصة في القضايا الحيوية لأخلقة هذا الذكاء، كقضايا الشفافية وإمكانية التتبع والرقابة البشرية (Hildebrandt, 2015). حيث يمكن أن تكون مختلف هذه التقاطعات عاملاً فاعلاً في الوصول إلى قواعد أكثر تميزاً وفاعلية لأخلقة هذا الذكاء، وكذلك حوكمته. كما نرى أنه يجب العمل لسد الفجوة القائمة بين طبيعة هذه البرمجيات والواقع الفعلي، ضمن محورين رئيسين: يتناول الأول ما يمكن وصفه بسد الفجوة القانونية عبر إعادة النظر في فروع القانون المختلفة لقياس الفجوة بين القواعد والتكنولوجيا، كما في قانون العقود، وقانون المسؤولية المدنية، وقانون المنافسة والمستهلك، وقانون الملكية الفكرية، وقانون البيانات الشخصية، وغيرها من القوانين ذات الصلة. إضافةً لسد الفجوة التقنية لهذا الذكاء، عبر تعزيز الشراكات المتبادلة بين مختلف العلوم المساهمة بشكل مباشر أو غير مباشر في هذه البرمجيات ضمن نهج متعدد التخصصات يجمع بين المحامين والقضاة وعلماء الرياضيات وعلماء الكمبيوتر بشكل خاص، وباقي المختصين في العلوم الاجتماعية والاقتصادية، بما يُمَكِنْ من دمج خصوصيات الذكاء الاصطناعي في معيار يستوعب الجميع ويراعي هواجس الجميع.

الخاتمة

وفق دراسة نقدية في الفلسفة والتأصيل، تناول البحث أحد جوانب أخلقة الذكاء الاصطناعي الأكثر تعقيداً والمرتبطة بتحديد "أين نحن؟" من أخلاقيات الذكاء الاصطناعي، وذلك ضمن سؤالين قوامها الاعتبار الأخلاقي والمقياس القانوني. في قضيةٍ تُعد اليوم واحدةً من القضايا الأكثر حضوراً في الفقه القانوني الغربي، والتي لا محالة ستكون موضع بحث قريب على الساحة الفقهية العربية. وقد عرضت الدراسة لذلك وفق مخطط بحثي حاول في قسمه الأول تحديد ماهية السلوك الأخلاقي الاصطناعي المُنتظر، بينما حاول قسمه الثاني توضيح المقياس أو المؤشر المعياري لهذه الأخلاق، بين الاعتبارين الاجتماعي أو القانوني.

منتهياً إلى خلاصة بحثية، يُمكن تمثيلها بحلقة معرفية تتكون من شِقّين كل منهما يفضي إلى الآخر ويخدم الآخر. حلقةٌ، شقها الأول مفاده: أن ضمان تصرف الذكاء الاصطناعي بأخلاقية يجب أن يكون مقدماً على المفهوم الصرف لأخلاق هذا الذكاء. وشقها الثاني، يؤكد: على أن تهذيب الذكاء الاصطناعي هو البوابة الموصلة لذكاء اصطناعي يتصرف بأخلاقية.

أولاً- أولوية التصرف بأخلاقية على التصرف الأخلاقي

إذ يؤكد البحث أن هناك اختلافاً واضحاً وجوهرياً بين التصرف الأخلاقي والتصرف بأخلاقية، فإنه يوصي بأن تركز الجهود والبرمجيات التي تتناول أخلاقيات الذكاء الاصطناعي على بلورة ذكاء اصطناعي يتصرف بأخلاقية. مشدداً على أن هذا الذكاء مُعرضٌ -وهو الواقع الفعلي- لتحديد الخيار الأخلاقي الأصح بين جملة خيارات أخلاقية جميعها صحيحة، فالإجابة الأخلاقية التي ننتظرها من الذكاء الاصطناعي ليست إجابة عن سؤال رياضي أو حسابي، أو حتى عن سؤال أخلاقي بسيط! بل عليه أن يكون مستعداً للإجابة عن أسئلة أخلاقية معقدة، بل وشديدة التعقيد والتركيب تقتضي الترجيح والاستنساب، مع ما يمكن أن يبرزه ذلك من إشكاليات، ببن المفهوم الإنساني للترجيح والاستنساب الذي يعتمد على الضمير والمشاعر وحتى العواطف، ونظيره الاصطناعي الذي يعتمد على آلية حسابية رقمية خوارزمية تفتقد لأي من الأحاسيس والمشاعر السابقة، والتي وإن بُرمجت خوارزمياً، تبقى مفتقدة حسياً.

وإذ يؤكد البحث أن هذا الفصل بين التقني والحسي هو حقيقة قائمة يصعب تلافيها ضمن المدى المنظور، فإنه يشدد ذلك يجب ألّا يكون عائقاً في سبيل الوصول لأفضل النتائج الخوارزمية الممكنة التي تضمن الوصول إلى القرار الخوارزمي ذي الصبغة الرقمية الأقرب للمفهوم الإنساني. مُوصياً بضرورة تعزيز أخلاقيات مطوري هذه البرمجيات وزرع المفاهيم الأخلاقية لديهم، ومن ثم تعزيز دائرة التشاركية في المفهوم التكويني للذكاء الاصطناعي، ضمن مفهوم التنشئة -الابتكار- الصحيحة والأخلاقية التي ستوجد لدينا في المحصلة ذكاءً اصطناعياً أخلاقياً مسؤولاً. وإن لم يكن هو بذاته موضع المساءلة أو الملاحقة(Chone-Grimaldi et Glaser, 2018, alerte 1. Courtois, 2016, p. 289. |الخطيب، 2020. ص. ص. 107-

151. عيسى، 2022، ص. ص. 210-403. بطيخ، 2021، ص. ص. 1513 - 1616. عبد ربه، 2021، ص. ص. 283 - 392. أحمد، 2020، ص. ص. 11-45.).

ثانياً- تحييد المعايير التقنية لمصلحة المعايير الإنسانية

إذ يؤكد البحث أنه لا يمكن الركون إلى المعايير التقنية المُستمدة من البيئة الاصطناعية الخالصة لهذا الذكاء في تحديد أخلاقيات هذا الذكاء، نظراً لانعدام هذه البيئة المؤهلة لبروز هكذا معايير، فإنه يشدد على ضرورة مراعاة المعايير الإنسانية في هذا الذكاء، مُوصياً بأن يكون ذلك وفق عملية تكيفية تأخذ بعين الاعتبار جميع هذه المعايير دون أي ترجيح أو تفضيل بينها. كونه لا يمكن تبني معيار واحد منفصل في تكوين هذه الأخلاقيات، ذلك أن أخلاقيات الذكاء الاصطناعي، عدا عن كونها استمراراً لأخلاقيات القائمين على هذا الذكاء، فهي في الفلسفة والتأصيل مرتبطة بالأخلاقيات ضمن مفهومها الإنساني والذي لا يساهم في تكونها عنصر واحد، بل هناك العشرات من العناصر الاجتماعية والاقتصادية والنفسية التي تساهم في بلورة هذه الأخلاقيات الإنسانية.

وإذ يؤكد البحث أن هذه الأخلاقيات الاصطناعية لا يمكن أن تُقاس بشكل منفرد بأي من المقاييس، سواء الاجتماعي أو القانوني منفرداً أو حتى الأخلاقي، إذ إن لكل منها عيوبه ومثالبه التي قد لا تتناسب مع الطبيعة الخاصة للذكاء الاصطناعي. فإنه يُشدد على أن مُختلف هذ التصورات وغيرها، هي التي تساهم في تكون هذه الأخلاقيات، وأن هذه العملية هي عملية تشاركية صرفة بين مختلف العلوم الإنسانية التي تنظر إلى الذكاء الاصطناعي كقيمة إنسانية مُضافة وليس كسلعة تجارية تنافسية، وذلك ضمن خلاصة مُفادها: "أن أخلقة الذكاء الاصطناعي هي مفتاح تهذيب الذكاء الاصطناعي وبأن تهذيب الذكاء الاصطناعي هي البوابة الموصلة لأخلقة الذكاء الاصطناعي". مُوصياً بضرورة احترام هذه المعادلة في أخلاقيات الذكاء الاصطناعي، بوصفها الدائرة المكتملة لحلقتين كلٌ منهما، هي البوابة والمَخرج للأخرى.

المراجع العربية:

زياد عبد الكريم القاضي. (2010). مقدمة في الذكاء الصناعي، الطبعة الأولى. دار صفاء للطباعة والنشر والتوزيع.

صفات سلامة، و خليل أبوقورة. (2014). تحديات عصر الروبوتات وأخلاقياته: دراسات استراتيجية، ، العدد 196، الطبعة الأولى. أبو ظبي: مركز الإمارات للدراسات والبحوث الاستراتيجية.

صلاح الفضلي. (2019). آلية عمل العقل عند الإنسان. القاهرة: عصر الكتب للنشر والتوزيع.

عابد فايد عبد الفتاح فايد. (2012). العدالة في القانون المدني، دراسة مقارنة بين القانون الفرنسي والمصري. مجلة البحوث القانونية والاقتصادية، كلية الحقوق، جامعة الإسكندرية، عدد خاص، ديسمبر، الصفحات 961-999.

عادل عبد النور. (2005). مدخل إلى عالم الذكاء الصناعي. منشورات جامعة الملك عبد العزيز.

عبد ربه المعداوي، و محمد أحمد مجاهد. (2021). المسئولية المدنية عن الروبوتات ذات الذكاء الاصطناعي: دراسة مقارنة. المجلة القانونية، جامعة القاهرة، المجلد التاسع العدد (2)، مصر، الصفحات 283-392.

فيصل محمد البحيري. (2015). أثر النهضة العلمية الحديثة على الفكر القانوني. الطبعة الأولى. مصر: مركز الدراسات العربية للنشر والتوزيع.

كريم موسى حسين. (2014). الذكاء الاصطناعي من منظور فلسفة العقل، العدد 11. مجلة الفلسفة. الجامعة المستنصرية، الصفحات 109-132.

ليث أثير يوسف. (2011). مشكلة الاستقراء والذكاء الصناعي. مجلة آداب المستنصرية. الجامعة المستنصرية. العدد 5، الصفحات 1-14.

محمد عبدالرازق، و وهبه سيد أحمد. (2020). المسؤولية المدنية عن أضرار الذكاء الاصطناعي: دراسة تحليلية. العدد 43، لبنان. مجلة جيل الأبحاث القانونية المعمقة، الصفحات 11-45.

محمد عبدالحفيظ المناصير. (2020). إشكالية الشخصية الإلكترونية القانونية للروبوت "دراسة تأصيلية تحليلية مقارنة في إطار التشريعين المدني العماني والأوربي. المجلة العربية للعلوم ونشر الأبحاث. المجلد (7)، العدد (1)، الصفحات 44-73.

محمد عرفان الخطيب. (2018). المركز القانوني للإنسآلة (Robots) قراءة في القواعد الأوربية في القانون المدني للإنسآلة (Robots) لعام 2017، "الإشكاليات والتحديات ... دراسة تأصيلية مقارنة". العدد الرابع. مجلة كلية القانون الكويتية العالمية. السنة السادسة، الصفحات 19-58.

محمد عرفان الخطيب. (2018). ضمانات الحق في العصر الرقمي، "من تبدل المفهوم .... لتبدل الحماية". قراءة في الموقف التشريعي الأوربي والفرنسي وإسقاط على الموقف التشريعي الكويتي. مجلة كلية القانون الكويتية العالمية. ملحق خاص، العدد الثالث، الجزء الأول، الصفحات 251-324.

محمد عرفان الخطيب. (2019). العدالة التنبؤية والعدالة القضائية، -الفرص والتحديات-، دراسة نقدية معمقة في الموقف الأنكلوسكسوني واللاتيني. المجلد 12، العدد 1. مجلة الحقوق والعلوم الإنسانية، جامعة زيان عاشور، الجفلة الجزائر، الصفحات 1-48.

محمد عرفان الخطيب. (2020). الذكاء الاصطناعي والقانون، دراسة نقدية مقارنة في التشريعين المدني الفرنسي والقطري في ضوء القواعد الأوربية في القانون المدني للإنسآلة لعام 2017 والسياسة الصناعية الأوربية للذكاء الاصطناعي والإنسآلات لعام 2019. ، وال. مجلة الدرسات القانونية، كلية الحقوق والعلوم السياسية، جامعة بيروت العربية، لبنان، المقال الرابع.

محمد عرفان الخطيب. (2020). المسؤولية المدنية والذكاء الاصطناعي، "إمكانية المساءلة؟!" دراسة تحليلية معمقة لقواعد المسؤولية المدنية في القانون المدني الفرنسي، العدد الأول. مجلة كلية القانون الكويتية العالمية. السنة الثامنة، الصفحات 107-151.

محمد عرفان الخطيب. (2020). سلسة الكتل "Blockchain"والعقود الذكية "Smart-Contracts" نموذجا" للتعايش بين القانون و العلوم المناظرة... الاقتصاد والرياضيات مثالا"، قراءة في الفلسفة والتأصيل. مجلة الدراسات القانونية، كلية الحقوق والعلوم السياسية جامعة بيروت العربية، لبنان، المقال الثالث.

محمد عرفان الخطيب. (2021). الذكاء الاصطناعي. الحاجة إلى التعريف "القانوني"؟! دراسة معمقة في الإطار الفلسفي للذكاء الاصطناعي من منظور قانوني مقارن". لبنان، المقال الأول. مجلة الدرسات القانونية، كلية الحقوق والعلوم السياسية، جامعة بيروت العربية.

مصطفى عيسى، و موسى أبو مندور . (2022). مدى كفاية القواعد العامة للمسئولية المدنية في تعويض أضرار الذكاء الاصطناعي، دراسة تحليلية تأصيلية مقارنة. مجلة حقوق دمياط للدراسات القانونية والاقتصادية، كلية الحقوق، جامعة دمياط، العدد الخامس، الصفحات 210-403.

مها رمضان محمد. (2021). المسؤولية المدنية عن أضرار أنظمة الذكاء الاصطناعي: دراسة تحليلية مقارنة، المجلة القانونية. المجلة القانونية، جامعة القاهرة،مصر، المجلد التاسع العدد الثاني، الصفحات 1513 - 1616.

ناصر شبكة. (2012). الذكاء الاصطناعي ومنطق تمثيل المعرفة. المجلد (1)، العدد (2). المجلة العربية الدولية للمعلوماتية، الصفحات 19-33.

المراجع الأجنبية:

Ananny, M. (2016). Toward an ethics of algorithms: Convening, observation, probability, and timeliness. Science, Technology, & Human Values, 41(1), Pp. 93-117.

Sadin, E. (2018). L'interview de Luc de Brabandere, L'intelligence artificielle ou l'enjeu du siècle. Anatomie d'un antihumanisme radical, éd. L'Échappée. Available online. .

AI Hleg. (2019a). High-level expert group on artificial intelligence.

Ananny, M. (2016). Toward an ethics of algorithms: Convening, observation, probability, and timeliness. Science, Technology, & Human Values, 41(1), Pp. 93-117.

Artificial intelligence and business models in the. (2020). : A systematic literature review, Journal of Business Research, Volume 121, Pp. 283-314.

Baron, B., & Musolesi, M. (2017). Interpretable machine learning for privacy-preserving pervasive systems. Pp. 1-10.

Bennett Moses, L. (2011). “Agents of Change.” 20 Griffith Law Review. Pp. 763-94. supra note 58.

Bensoussan, A. (2013). Plaidoyer pour un droit des robots : de la personne morale à la personne robot, La Lettre des juristes d'affaires, 23 oct, n° 1134.

Białek, M., & De Neys, W. (2016). "Conflict detection during moral decision-making: evidence for deontic reasoners' utilitarian sensitivity". Journal of Cognitive Psychology. 28 (5): 631-639.(2016-03-02). .

Blin-Franchomme, M.-P. (2021). Le défi d'une IA inclusive et responsable. Droit social p. 100.

Bloom, P. (2011). "Family, community, trolley problems, and the crisis in moral psychology". The Yale Review. 99 (2). Pp. 26-43.

Bolte, J., Bar, A., Lipinski, D., & Fingscheidt, A. (2019). Towards corner case detection for autonomous driving. In 2019 IEEE Intelligent Vehicles Symposium, Pp.438-445.

Bostrom, N., & Yudkowsky, E. (2014). The ethics of artificial intelligence. The Cambridge Handbook of Artificial Intelligence.

Bourcier, D. (2001). De l'intelligence artificielle à la personne virtuelle : émergence d'une entité juridique ? Droit et société, /3, n° 49, Pp. 847-871.

Bourcier, D. (2018). Le droit va-t-il disparaître dans les algorithmes ? LPA, n° 223, p. 8.

Brady, W., Wills, A., Jost, T., & Tucker, J. (2017). Emotion shapes the diffusion of moralized content in social networks. Proc Natl Acad Sci USA, 114(28), Pp. 7313-7318.

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In Sorelle and Wilson 2018. Pp. 1-15.

Cadre pour les aspects éthiques de l’intelligence . (2020). de la robotique et des technologies connexes. P9 TA 0275.

Castets-Renard, C. (2021). Quelle politique européenne de l'intelligence artificielle ? RTD Eur. p. 297.

Chassagnard-Pinet, S. (2017). Les usages des algorithmes en droit : prédire ou dire le droit ? Dalloz IP/IT. P. 495.

Chatila, R. (2016). Intelligence artificielle et robotique : un état des lieux en perspective avec le droit, Dalloz IP/IT, p. 284.

Chone-Grimald, A.-S., & Glaser, P. (2018). Responsabilité du fait du robot doué d’intelligence artificielle : faut-il créer une personnalité robotique ? Contrats Concurrence Consommation, n° 1, alerte 1.

CLIP . (2022). CLIP experiments revealed that images of Black people were misclassified as nonhuman more than twice as frequently as other races. Artificial Intelligence Index Report 2022.

Conseil de l'Europe. (2019). Charte éthique européenne d'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement, , févr.

Courtois, G. (2016). Robots intelligents et responsabilité : quels régimes, quelles perspectives ? D, p. 289.

Crawford, K., & Joler, V. (2018). Anatomy of an AI system. Available online. Kate Crawford and Vladan Joler, "Anatomy of an AI System: The Amazon Echo as An Anatomical Map of Human Labor, Data and Planetary Resources," AI Now Institute and Share Lab, Sept.

De Bruin, B., & Floridi, L. (2017). The ethics of cloud computing. Science and Engineering Ethics, 23(1), Pp. 21-39.

Dhar, P. (2020). The carbon impact of artificial intelligence. Nat Mach Intell 2, Pp. 423-425.

Dignum, V. (2019). Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way, Cham: Springer International Publishing.

Drouot, G. (2020). Droit, algorithmes et anarchie. Recueil Dalloz. p. 35.

Emmanuel, D., Mamed, D., Maloo, S., & Degannes, J. (2012). "Stress alters personal moral decision making". Psychoneuroendocrinology. 37 (4): 491-498.

Ethics Guidelines for Trustworthy. (2020). Ethics Guidelines for Trustworthy Available online March.

Favoreu, L. (2002). Résurgence de la notion de déni de justice et droit au juge, in Mélanges J. Waline, Dalloz, p. 513.

Floridi, L. (2018). Soft ethics, the governance of the digital and the General Data Protection Regulation. Phil. Trans. R. Soc. A 376, 20180081.

Forest, D. (2017). La régulation des algorithmes, entre éthique et droit, RLDI, n° 137, p. 38 s.

Foss-Solbrekk, K. (2021). Three routes to protecting AI systems and their algorithms under IP law: The good, the bad and the ugly, Journal of Intellectual Property Law & Practice, Vol 16, Issue 3, Pp. 247-258.

Frison-Roche, M.-A. (2003). "Déni de justice et interprétation de la loi par le juge", J-Cl droit civil, art. 4.

Gaudemet, Y. (2018). La justice à l'heure des algorithmes, RDP, p. 651.

Gebru, T., Morgenstern, J., Vecchione, B., Wallach, H., Daumeé, H., Crawford, K., & Vaughan, J. (2018). Datasheets for datasets. Pp. 1-17.

Glastra van, L. (1958). 'Rules and Commands', Mind LXVII, no. 268. Pp. 1-9.

Godefroy, L. (2018). Le code algorithmique au service du droit, D, p. 734.

Godefroy, L. (2020). Éthique et droit de l'intelligence artificielle, osmose ou symbiose ? D. 231.

Groupe d'experts indépendants de haut niveau sur l. (2019). Lignes directrices en matière d'éthique pour une IA digne de confiance, avr. Available online.

Groupe d'experts indépendants de haut niveau sur l. (2019). Lignes directrices en matière d'éthique pour une IA digne de confiance, Publications de l'UE.

Hagendorff, T. (2020). "The Ethics of AI Ethics: An Evaluation of Guidelines." Minds and Machines. Pp. 99-120.

Hagendorff1, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. Minds and Machines 30. Pp. 99-120. P. 99. P. 105.

Hildebrandt, M. (2015). Smart Technologies and the End (s) of Law: Novel Entanglements of Law and Technology, Cheltenham: Edward Elgar Publishing.

Hildebrandt, M. (2017). The Force of Law and the Force of Technology. In The Routledge International. Handbook of Technology, Crime and Justice, ed. M.R.P. McGuire and Holt. Routledge, Pp. 579-608.

Howard, P., & Kollanyi, B. (2016). Bots, #StrongerIn, and #Brexit: Computational propaganda during the UK-EU Referendum. Pp. 1-6.

Irani, L. (2015). The cultural work of microwork. New Media & Society, 17(5), Pp. 720-739.

Jarvis Thomson , J. (1985). "The Trolley Problem" (PDF). Yale Law Journal. 94 (6). Pp. 1395-1415.

Jobin, A., Marcello, I., & Effy, V. (2019). “The Global Landscape of AI Ethics Guidelines.”. 1 Nature Machine Intelligence Pp. 389-99.

Julia, A., Jeff, L., Surya, M., & Lauren, K. (2016). ProPublica, there’s software used across the country to predict future criminals. And it’s biased against blacks. May 23. Available online.

Kitchin, R. (2017). Thinking critically about and researching algorithms. Information, Communication & Society, 20(1), Pp. 14-29.

Kitchin, R., & Dodge, M. (2011). Code/space: Software and everyday life. Cambridge: The MIT Press.

Koulu, R. (2020). “Human Control over Automation: EU Policy and AI Ethics.” 12 European Journal of Legal Studies. Pp. 9-46. supra note 37.

Kritikos, M. (2019). Artificial Intelligence ante portas: Legal & ethical reflections. STOA | Panel for the Future of Science and Technology. EPRS | European Parliamentary Research Service. P. 634 & P. 427.

Kroll, J. (2018). The fallacy of inscrutability. Phil. Trans. R. Soc. A 376, 20180084. p. 11.

Langlois, L., intervention, S., & Summit , A. (2019). La régulation de l'IA par l'éthique : une alliance pour une plus grande vigilance, 20 nov.

Larrière, S. (2017). Confier le droit à l’intelligence artificielle : le droit dans le mur ? Revue Lamy droit de l'immatériel, n° 134, Pp. 38-40.

Larsson, S. (2019). “The Socio-Legal Relevance of Artificial Intelligence.” 103 Droit et Société. Pp. 573-93. supra note 16, in terms of a “mirror for social structures.”.

Larsson, S. (2020). On the Governance of Artificial Intelligence through Ethics Guidelines. Asian Journal of Law and Society, 7, Pp. 437-451.

Larsson, S. (2020). On the governance of artificial intelligence through ethics guidelines. Journal of Law and Society, 1, 23.

Le Ny , J.-F. (1993). Intelligence Naturelle et Intelligence Artificielle. Paris, Presses Univ., de France, p. 12.

Leiser, M., & Murray, A. (2016). The role of non-state actors and institutions in the governance of new and emerging digital technologies. In The oxford handbook of law, regulation and technology, Eds. R. Brownsword, E. Scotford, K. Yeung, & O. U. Press.

Lessig, L. (1999). Code is Law, The Industry Standard, 18.

Lessig, L. (2000). Code is Law. On liberty in cyberspace, Harvard Magazine.

Lignes directrices en matière d'éthique pour une I. (2018). Groupe d'experts de haut niveau sur l'IA constitué par la Commission européenne.

Lim, H.-S., & Taeihagh, A. (2019). "Algorithmic Decision-Making in AVs: Understanding Ethical and Technical Concerns for Smart Cities". Sustainability. 11 (20): 5791.

Loiseau, G. (2018). La personnalité juridique des robots : une monstruosité juridique, Libres propos, JCP. G, vol. 22, Pp. 1039-1042.

Martha, N. (2015). The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy, Cambridge: Cambridge University Press, 2001, 95. parrhesia 23. Pp. 122-43.

Meirieu, P. (1992). Éduquer : un métier impossible ? ou « Éthique et pédagogie », . Pédagogie Collégiale Septembre, Vol. 6 n° 1. Conférence à Montréal 27 mai 1992.

Meirieu, P. (1992). Éduquer : un métier impossible ? ou « Éthique et pédagogie », Conférence à Montréal 27 mai 1992. Pédagogie Collégiale Septembre, Vol. 6 n° 1.

Mendoza-Caminade, A. (2016). Le droit confronté à l'intelligence artificielle des robots : vers l'émergence de nouveaux concepts juridiques ? D, p. 445.

Meneceur, Y. (2018). Quel avenir pour la justice prédictive ? Enjeux et limites des algorithmes d'anticipation des décisions de justice, JCP G, p. 316.

Michael, A., & Susan Leigh, A. (2007). Machine Ethics: Creating an Ethical Intelligent Agent, Artificial Intelligence Magazine, 28:4, Winter, Pp. 15-26.

Mihalis , K. (2018). What if algorithms could abide by ethical principles? EPRS | European Parliamentary Research Service. Scientific Foresight Unit (STOA) PE 624.267. November.

Mihalis , K. (2019). Artificial Intelligence ante portas : Legal & ethical reflections. STOA | Panel for the Future of Science and Technology. EPRS | European Parliamentary Research Service. PE 634.427.

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. Nature Machine Intelligence, 1(11), Pp. 501-507.

Mittelstadt, B., Russell, C., & Wachter, S. (2019). Explaining explanations in AI. In Proceedings of the conference on fairness, accountability, and transparency. FAT 19. Pp. 1-10.

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2019). From what to how. An overview of AI ethics tools, methods and research to translate principles into practices. Pp. 1-21.

Oliver, W.-H. (1997). 'The Path of the Law', Harvard Law Review 110. Pp. 991-1009.

Oster, J. (2021). Code is code and law is law. The law of digitalization and the digitalization of law, International Journal of Law and Information Technology, Vol. 29, Issue 2, Pp. 101-117.

Pei, K., Cao, Y., Yang, J., & Jana, S. (2017). Towards practical verification of machine learning: The case of computer vision systems. Vol. 1712. 01785. Available online.

Pekka, A., Bauer, W., Bergmann, U., Bieliková, M., Bonefeld-Dahl, C., & Bonnet, Y. (2018). On the Governance of Artificial Intelligence through Ethics Guidelines. Asian Journal of Law and Society, 7 (2020), pp. 437-451.

Philippa, F. (1978). "The Problem of Abortion and the Doctrine of the Double Effect" in Virtues and Vices. Oxford: Basil Blackwell. Originally appeared in the Oxford Review, Number 5, 1967.

Quenillet, M. (1994). Droit et intelligence artificielle : mythes, limites et réalités, Petites affiches, n° 66, p. 11.

Rahwan, I. (2018). Society-in-the-loop: Programming the algorithmic social contract. Ethics and Information Technology, 20(1), Pp. 5-14.

Règl, U. (2016). 679 du Parlement européen et du Conseil du 27 avr. 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données (RGPD), JOUE L 119/1, 4 mai.

Résolution du Parlement européen. (2020). contenant des recommandations à la Commission concernant un cadre pour les aspects éthiques de l'intelligence artificielle, de la robotique et des technologies connexes (2020/2012(INL).

Résolution du Parlement européen. (2020). contenant des recommandations à la Commission sur un régime de responsabilité civile pour l'intelligence artificielle (2020/2014(INL), P9 TA (2020)0276.

Ricoeur, P. (1988). Avant la loi morale, l'éthique. , Les enjeux, Pp. 42-45. Encyclopedia Universalis.

Ricoeur, P. (1988). Avant la loi morale, l'éthique. Encyclopedia Universalis, Les enjeux, Pp. 42-45.

Rouvière, F. (2017). La justice prédictive, version moderne de la boule de cristal, RTD civ. P. 527.

Rouvière, F. (2020). L'intelligence artificielle au risque du mythe. RTD Civ, p. 990.

Samuel, M., & Henri, I. (2000). Ethique ou déontologie : quelles différences pour quelles conséquences managériales ? L'analyse comparative de 30 codes d'éthique et de déontologie, , 05, Montpellier. IXème Conférence AIMS.

Samuel, M., & Henri, I. (2000). Ethique ou déontologie : quelles différences pour quelles conséquences managériales ? L'analyse comparative de 30 codes d'éthique et de déontologie, IXème Conférence AIMS, 05, Montpellier.

Silberman, M., Tomlinson, B., LaPlante, R., Ross, J., Irani, L., & Zaldivar, A. (2018). Responsible research with crowds. Communications of the ACM, 61(3), Pp. 39-41.

Simpson, H., & Hale, S. (1969). Pupillary changes during a decision-making task. Mot. Skills 29, Pp. 495-498.

Singer, P. (1972). "Famine, Affluence, and Morality". Philosophy & Public Affairs. 1 (3): 229-243.

Skulmowski, A., Bunge, A., Kaspar, K., & Pipa, G. (2014). Forced-choice decision-making in modified trolley dilemma situations: a virtual reality and eye tracking study. Front. Behav. Neurosci. 8:426.

Ştefan, T.-M. (2020). Ethics in Artificial Intelligence. International Journal of User-System Interaction 13(3), Pp. 136-148.

Sütfeld, L., Gast, R., König, P., & Pipa, G. (2017). "Using Virtual Reality to Assess Ethical Decisions in Road Traffic Scenarios: Applicability of Value-of-Life-Based Models and Influences of Time Pressure". Frontiers in Behavioral Neurosc.

Taeihagh, A., Ramesh, M., & Howlett, M. (2021). Assessing the regulatory challenges of emerging disruptive technologies. In Regulation & Governance. Available online.

The ethics of artificial intelligence. (2020). : Issues and initiatives. STOA | Panel for the Future of Science and Technology. EPRS | European Parliamentary Research Service. PE 634.452.

The European Commission's high-level expert group . (2018). Ethics guidelines for trustworthy ai. Working Document for stakeholders' consultation. Brussels. Pp. 1-37.

Unger, P. (1996). Living High and Letting Die. Oxford: Oxford University Press.

Veale, M. (2020). "A Critical Take on the Policy Recommendations of the EU High-Level Expert Group on Artificial Intelligence." European Journal of Risk Regulation. Pp. 1-8. Supra note 37.

Veglis, A. (2014). Moderation techniques for social media content. In D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, A. Kobsa, F. Mattern, J. C. Mitchell, et al. (Eds.), Social computing and social media. Pp. 137-148.

Wagner, B. (2018). Ethics as an escape from regulation: From « Ethics-Washing » To « Ethics-Shopping »? in Being Profiled, Cogitas Ergo Sum: 10 Years of Profiling the European Citizen, E. Bayamlioglu, I. Baraliuc, L. Janssens, M. Hildebrandt, Amsterdam University Press.

Wang, P. (2019). On Defining Artificial Intelligence. Journal of Artificial General Intelligence, 10 (2). Pp. 1-37.

Yampolskiy, R.-V. (2020). On Defining Differences Between Intelligence and Artificial Intelligence. Journal of Artificial General Intelligence 11(2), Pp. 68-70.

Yampolskiy, R.-V., & Fox, J. (2012). Artificial general intelligence and the human mental model. In A-H. Eden, J-H. Soraker, & E. Steinhart (Eds.), Singularity hypotheses, Pp. 129-145.

Yeung, K., & Martin, L. (2019). Algorithmic Regulation, Oxford: Oxford University Press.

Zhang, B., & Dafoe, A. (2020). US public opinion on the governance of artificial intelligence. In Proceedings of the AAAI/ACM Conference on AI, Ethics and Society, pp.187-193.