المعايير الأخلاقية لتصميم وتطوير تطبيقات الذكاء الاصطناعى فى العصر الرقمى

■ أ.د/ محمد فوزى والى

أستاذ تكنولوجيا التعليم بكلية التربية

عميد كلية الحاسبات والمعلومات بجامعة دمنهور السابق

■ ملخص البحث:

يشهد مجال الذكاء الاصطناعى تطورًا متسارعًا مع تزايد تطبيقاته فى مختلف جوانب الحياة، مما يطرح تساؤلات جوهرية ومتجددة حول كيفية تطوير واستخدام هذه التطبيقات بطريقة أخلاقية ومسئولة ومفيدة للمجتمع ككل. وتُشير أخلاقيات التعامل مع تطبيقات الذكاء الاصطناعى إلى مجموعة المبادئ والقيم التى توجه تطوير واستخدام هذه التطبيقات بشكل مسئول، بما يضمن تحقيق أقصى استفادة منها مع تقليل كافة المخاطر المحتملة، ولقد استهدف البحث الحالى تقصى أهم المعايير الأخلاقية لتصميم وتطوير تطبيقات الذكاء الاصطناعى فى العصر الرقمي، وتم التوصل إلى قائمة بالمعايير الأخلاقية لتصميم وتطوير تلك التطبيقات؛ وتضمنت هذه القائمة خمسة عشر معيارًا أهمها: المسئولية، والمساءلة، والشفافية، والعدالة، والخصوصية، والاستدامة البيئية، ولقد أوصى البحث بضرورة مراعاة هذه المعايير ومؤشراتها الفرعية عند تصميم وتطوير تطبيقات الذكاء الاصطناعى حيث تساعد هذه المعايير فى توجيه عملية تصميم وتطوير تطبيقات الذكاء الاصطناعى لضمان تحقيق أقصى استفادة منها مع تقليل المخاطر المحتملة، فضلا عن اقتراح البحث ضرورة إجراء مزيد من البحوث لتناول الآثار التى ستترتب على مراعاة تلك المعايير الأخلاقية من وجهة نظر كل من المُطورين والمُستخدمين.

■ Abstract:

The field of artificial intelligence (AI) is witnessing rapid advancement, accompanied by a growing range of applications across various domains of life. This evolution raises essential and continually evolving questions regarding how to develop and utilize AI applications in an ethical, responsible, and socially beneficial manner. The ethics of AI applications refer to a set of principles and values that guide the responsible development and deployment of these technologies, aiming to maximize their benefits while minimizing potential risks.

This study seeks to investigate the most critical ethical standards that should govern the design and development of AI applications in the digital era. The research culminated in a proposed list of ethical standards for designing and developing such applications. This list comprises fifteen key standards, including responsibility, accountability, transparency, fairness, privacy, and environmental sustainability.

The study recommends the integration of these standards and their respective indicators into the design and development processes of AI applications. These ethical guidelines serve as a compass to ensure optimal benefits from AI while mitigating associated risks. Furthermore, the research advocates for additional studies to explore the implications of applying these ethical standards from the perspectives of both developers and users.

الكلمات المفتاحية:

المعايير الأخلاقية، تصميم وتطوير تطبيقات الذكاء الاصطناعي، استراتيجية الذكاء الاصطناعى، استراتيجية مكافحة الفساد.

مقدمة البحث:

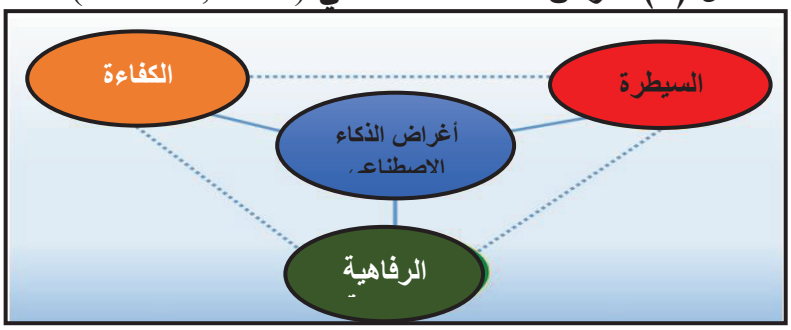
لقد أصبح الذكاء الاصطناعي Artificial Intelligence (AI) جزءًا لا يتجزأ من حياتنا اليومية، ويتم استخدامه على نطاق واسع، ويتطور هذا المجال بصورة ملحوظة، ولقد عرف كل من "لوسين؛ وبينجو؛ وهينتون"(LeCun, Bengio & Hinton, (2015 الذكاء الاصطناعي على أنه مجموعة من التقنيات التي تسمح للحواسيب بفهم العالم من حولها والتفاعل معه بطريقة ذكية، من خلال معالجة البيانات، واتخاذ الإجراءات المناسبة، وتطوير نماذج تنبؤية.

كما عرفه "ستون" وآخرون Stone et al (2016) على أنه مجال واسع يهدف إلى تطوير أنظمة قادرة على محاكاة القدرات المعرفية البشرية، مثل: التعلم والاستنتاج وحل المشكلات واتخاذ القرارات والإبداع.

وأشار أحمد العربي (٢٠٢١) إلى أن الذكاء الاصطناعي مجال واسع يشمل عديدًا من التخصصات، مثل: تعلم الآلة، ومعالجة اللغات الطبيعية، والرؤية الحاسوبية، والروبوتات، ويهدف إلى تطوير أنظمة ذكية قادرة على العمل بشكل مستقل واتخاذ القرارات، كما أوضح كل من "روسل" و"نورفيج" Russell & Norvig (2021) أن الذكاء الاصطناعي فرع من فروع علوم الحاسوب يركز على تصميم وبناء أنظمة ذكية قادرة على العمل بشكل مستقل واتخاذ القرارات في مواقف معقدة وغير مؤكدة. وخلاصة القول: إن الذكاء الاصطناعي يعد أحد أهم التوجهات الحديثة، حيث إنه أحد علوم الحاسب الآلي الحديثة التي تهدف إلى فهم طبيعة الذكاء الإنساني، ومتابعة انتباهه واستجاباته، ثم القدرة على محاكاة السلوك الإنساني ورد الفعل المتسم بالذكاء لتقديم نصيحة أو حل مسألة أو المساعدة في اتخاذ قرار Sreenu, (2023).

ويمتلك الذكاء الاصطناعي (AI) إمكانيات هائلة لتغيير المجتمع. وتؤدي التطبيقات الواسعة للأنظمة الذكية إلى تحقيق أرباح اقتصادية، مما يؤدي إلى وجود مجموعة من التحديات الأخلاقية المصاحبة للذكاء الاصطناعي. ويتفق معظم الباحثين والمطورين على ضرورة تطوير الذكاء الاصطناعي بطريقة موثوقة تفيد المجتمع ككل ويعرض الشكل التالي أغراض استخدام تطبيقات الذكاء الاصطناعي (Berendt, 2019; Cath et al., 2018; European Commission, 2020; Floridi et al., 2018; Jobin et al., 2019).

شكل (١) أغراض الذكاء الاصطناعي (Berendt, 2019)



ويتطلب تصميم وتطوير تطبيقات الذكاء الاصطناعي ضرورة التغلب على التحديات الأخلاقية. وعلى الرغم من توافر عديد من الدراسات التي تناولت تطوير تطبيقات الذكاء الاصطناعي، إلا أن هناك نقصًا كبيرًا في الدراسات المرتبطة بتحديد المعايير الأخلاقية لاستخدام هذه التطبيقات بشكل دقيق ومحدد (Shin & Park, 2019).

ويعد التوازن بين المعايير الأخلاقية عند تصميم تطبيقات الذكاء الاصطناعي أمرًا بالغ الأهمية، لأن تحقيقها جميعًا في آنٍ واحد غالبًا ما يكون صعبًا أو حتى مستحيلًا أحيانًا. وعلى الرغم من أن تطبيقات الذكاء الاصطناعي تستهدف في الأساس تخفيف الأعباء على المستخدمين وتسهيل تنفيذهم لمختلف العمليات، إلا أن إتاحة البيانات الشخصية للأفراد قد تثير مخاوف كبيرة من بينها مثلا انتهاك الخصوصية (Marwick& Lewis,, 2017; Shin & Park, 2019).

ولقد أشار فريق من الخبراء معني بالذكاء الاصطناعي (AI HLEG) تابع للمفوضية الأوروبية إلى ضرورة توافر سبعة متطلبات أساسية لضمان توافر تطبيقات ذكاء اصطناعي موثوق بها وتمثلت هذه المتطلبات في (Shneiderman, 2020):

- الوكالة البشرية والرقابة
- المتانة التقنية والسلامة
- الخصوصية وحوكمة البيانات
- الشفافية
- التنوع وعدم التمييز والعدالة
- الرفاهية الاجتماعية والبيئية
- المساءلة

كما أوصت منظمة التعاون الاقتصادي والتنمية (OECD) بضرورة مراعاة خمسة مبادئ أخلاقية عند الشروع في تصميم تطبيقات الذكاء الاصطناعي وجاءت هذه المبادئ على النحو التالي (OECD, 2021):

- النمو الشامل والتنمية المستدامة والرفاهية
- القيم الإنسانية والعدالة
- الشفافية وقابلية التفسير
- المتانة والأمان والسلامة
- المساءلة

وقد قام كل من "جوبين" وآخرون Jobin et al (2019) بمراجعة ٨٤ بحثًا من جميع أنحاء العالم حول أخلاقيات الذكاء الاصطناعي، وحدد المؤلفون (١١) مبدأً أخلاقيًا شاملاً، ووجدوا أن خمسةً منها هي الشفافية، العدالة والإنصاف، عدم الإضرار، المسئولية، والخصوصية، تم ذكرها في أكثر من نصف البحوث التي تم تحليلها. كما تم العثور على مبادئ المنفعة (Beneficence) والحرية والاستقلالية في عدد (٤١) بحثًا على التوالي. أما المبادئ الأخلاقية الأخرى مثل: الثقة، والاستدامة، والكرامة، والتضامن، فقد تم ذكرها فيما يقرب من ثلث البحوث.

وعلى الرغم من أن جميع المبادئ الأخلاقية المذكورة تبدو مرغوبة من حيث المبدأ، فإن تطبيقها في الواقع العملي قد يواجه تحديات كبيرة. والسبب في ذلك هو أنه عند تصميم تطبيقات ذكاء اصطناعي، غالبًا ما يكون من الصعب تحقيق جميع الجوانب الأخلاقية المختلفة في آنٍ واحد (Köbis et al., 2021) .

وتشكل الأخلاقيات والقوانين عنصرين رئيسيين في الإطار التنظيمي الذي يحكم تطوير واستخدام تطبيقات الذكاء الاصطناعي. فالأخلاقيات توفر المبادئ التوجيهية التي تضمن أن تكون التطبيقات الذكية متوافقة مع القيم الإنسانية، مثل: العدالة، والشفافية، والمسئولية. بينما تعمل القوانين على وضع قواعد وإجراءات ملزمة قانونًا لتنظيم استخدام هذه التقنية والحد من أي انتهاكات محتملة، مثل انتهاك الخصوصية أو التحيز، وتجدر الإشارة إلى أن التكامل بين الجوانب الأخلاقية والتشريعية يساهم في تطوير ذكاء اصطناعي مسئول يخدم المجتمع بشكل عادل وآمن (Bryson, 2018).

ومن الجدير بالذكر أن أخلاقيات التعامل مع تطبيقات الذكاء الاصطناعي ليست مسئولية المطورين والباحثين فحسب، بل هي مسئولية مشتركة تشمل أيضا كلًّا من المستخدمين والشركات والحكومات، والمنظمات، والمجتمع ككل. فمن خلال التعاون والالتزام بالمبادئ الأخلاقية يمكننا ضمان أن يكون الذكاء الاصطناعي قوة إيجابية تُسهم في بناء مستقبل أفضل للجميع.

ولذلك؛ تسعى الحكومات والمؤسسات الأكاديمية والمنظمات التقنية إلى وضع سياسات وإرشادات أخلاقية لضمان تصميم وتطوير واستخدام الذكاء الاصطناعي بشكل آمن ومسئول. فقد أصدرت عديد من المنظمات مثل: الاتحاد الأوروبي ومنظمة التعاون الاقتصادي والتنمية (OECD) أطرًا توجيهية لتطوير أنظمة ذكاء اصطناعي يراعي الاعتبارات الأخلاقية (Jobin et al., 2019). كما أصدرت جمهورية مصر العربية الاستراتيجية الوطنية للذكاء الاصطناعي كخطوة رائدة لتعظيم الاستفادة من تطبيقات الذكاء الاصطناعي، ومع ذلك، لا تزال هناك تحديات في التطبيق العملي لهذه المبادئ، مما يتطلب تضافر الجهود بين الباحثين وصناع القرار والمطورين لضمان توافق أنظمة الذكاء الاصطناعي مع القيم الإنسانية الأساسية (Mittelstadt et al., 2016).

مشكلة البحث:

تعاني بعض التطبيقات الخاصة بالذكاء الاصطناعي في العصر الرقمي من عدم مراعاة البعد الأخلاقي، الأمر الذي قد يترتب عليه عديد من المشكلات التي يُمكن أن تضر المجتمع من بينها تعزيز التحيز والتمييز ضد بعض الفئات الاجتماعية، فضلًا عن انتهاك الخصوصية وسوء استخدام البيانات حيث قد يتم جمع بيانات المستخدمين دون إذن مسبق مما قد يؤدي لمخاطر مثل: السرقة الرقمية، والابتزاز، وانتهاك الخصوصية الشخصية، كما يمكن استغلال تطبيقات الذكاء الاصطناعي في التضليل الإعلامي، مثل: إنشاء أخبار مزيفة أو مقاطع فيديو مزورة، فضلا عن تقليل التفاعل البشري المباشر، مما يزيد من العزلة الاجتماعية ويمكن تلخيص أهم الآثار التي ترتبت على الوضع الحالي لاستخدام تطبيقات الذكاء الاصطناعي في النقاط التالية:

- **التحيز في البيانات:** يمكن أن تزيد التحيزات الموجودة في البيانات من احتمالية تحيز نماذج الذكاء الاصطناعي، مما يؤدي إلى نتائج غير عادلة، مثل: التحيزات العرقية أو الاجتماعية والاقتصادية.
- **الخصوصية والأمان:** قد يؤدي التعامل غير السليم مع البيانات الحساسة إلى انتهاكات للخصوصية أو تسريبات للبيانات.
- **جودة البيانات ونزاهتها:** يمكن أن تؤدي البيانات ذات الجودة المنخفضة أو غير المكتملة أو القديمة إلى تنبؤات غير موثوقة للنموذج، مما يؤدي إلى قرارات غير دقيقة.
- **التحيز في الخوارزميات:** إلى جانب البيانات، قد تؤدي تصميمات الخوارزميات إلى تحيزات إذا لم يتم تطويرها بعناية، مما يؤدي إلى نتائج غير عادلة أو تمييزية.
- **التحيز في الأتمتة:** قد يعتمد المستخدمون بشكل مفرط على توصيات الذكاء الاصطناعي، بافتراض أنها أكثر دقة مما هي عليه في الواقع.
- **فشل الأنظمة:** يمكن لأنظمة الذكاء الاصطناعي أن تفشل أو تتصرف بشكل غير متوقع في البيئات الديناميكية، خاصة عند مواجهة سيناريوهات غير متوقعة.
- **مشكلات القابلية للتوسع:** مع نمو أنظمة الذكاء الاصطناعي، قد تصبح صيانتها والمحافظة على كفاءتها وجودتها عبر بيئات مختلفة أمرًا أكثر تعقيدًا.
- **الهجمات العدائية:** قد تكون أنظمة الذكاء الاصطناعي، خاصة في مجالات مثل: الأمن السيبراني، عرضة للهجمات حيث يتم التلاعب بالمدخلات لإنتاج مخرجات مغلوطة.
- **العدالة والتمييز:** قد تزيد تطبيقات الذكاء الاصطناعي من عدم المساواة أو تؤدي إلى أشكال جديدة من التمييز إذا لم يتم إدارتها بعناية.
- **الشفافية والمساءلة:** يمكن أن يؤدي نقص الشفافية في عمليات اتخاذ القرار في الذكاء الاصطناعي إلى صعوبة تحديد المسئولية عند حدوث أخطاء.

- **التزييف والمعلومات المضللة:** يمكن استخدام المحتوى الذي يتم إنشاؤه بواسطة الذكاء الاصطناعي، مثل التزييف لنشر معلومات مضللة بطريقة ضارة، مما يؤثر على المصادر العامة للمعلومات.
- **قضايا الملكية الفكرية:** يمكن أن يؤدي سوء استخدام البيانات أو المواد المحمية بحقوق النشر إلى نزاعات حول الملكية الفكرية.

وبناءً عليه أصبحنا في حاجة لتحديد مجموعة من المعايير الأخلاقية الواجب مراعاتها أثناء تصميم وتطوير تطبيقات الذكاء الاصطناعي. ويرجع ذلك إلى الطبيعة المتطورة لهذه التقنية وتعقيداتها المتزايدة، مما استدعى وضع أسس منهجية تضمن توافق أنظمة الذكاء الاصطناعي مع المبادئ الأخلاقية، مثل: الشفافية، والعدالة، والمسئولية. وعليه سعى هذا البحث إلى استقصاء وتحديد قائمة بالمعايير الأخلاقية التي يمكن اعتمادها لضمان تطوير أنظمة ذكاء اصطناعي تعزز القيم الإنسانية وتحُد من المخاطر المحتملة المرتبطة باستخدامه.

أهداف البحث:

يسعى البحث الحالي إلى:

١-تحديد أهم المعايير الأخلاقية الخاصة بتصميم وتطوير تطبيقات الذكاء الاصطناعي في مختلف المجالات. ويشمل ذلك قضايا مثل: الخصوصية، والعدالة، والتحيز، والمسئولية القانونية، والتأثيرات الاجتماعية لهذه التقنية.

٢-دراسة الأساليب والاستراتيجيات التي يمكن اعتمادها لضمان التعامل مع هذه القضايا بشكل مسئول، بما يحقق التوازن بين التطور التكنولوجي السريع وحماية القيم الإنسانية الأساسية*.

٣-تقديم حلول وتوصيات تسهم في توجيه تطوير واستخدام تطبيقات الذكاء الاصطناعي بطريقة تعزز التنمية المستدامة وتقلل من مخاطر الفساد المحتم.

٤- تحليل الموقف الحالي الخاص بأخلاقيات تصميم وتطوير تطبيقات الذكاء الاصطناعي فيما يرتبط حجم المشكلة، أسباب المشكلة، آثار المشكلة، السياسات السابقة، الأطراف الفاعلة سلبًا أو إيجابًا، المؤشرات الإيجابية والسلبية بصدد الموقف الحالي.

فروض البحث:

يسعى البحث الحالي إلى دراسة الفروض التالية:

١. توجد علاقة ارتباطية بين مراعاة أخلاقيات تصميم وتطوير تطبيقات الذكاء الاصطناعي وتقليل المشكلات والتحديات الخاصة بالاستخدام غير المسئول لهذه التطبيقات.

---

* القيم الإنسانية الأساسية: هي مجموعة من المبادئ الأخلاقية التي توجه سلوك الأفراد والمجتمعات، وتشمل قيمًا مثل الكرامة الإنسانية، والعدالة، والمساواة، والحرية، والتضامن، والسلام. هذه القيم تعتبر أساساً للتعايش السلمي والتعاون بين الناس، وتسهم في بناء مجتمعات مزدهرة يسودها الاحترام المتبادل والتقدير للتنوع (Schwartz, 1992).

٢. توجد علاقة ارتباطية بين وجود معايير أخلاقية واضحة لتصميم وتطوير واستخدام أنظمة الذكاء الاصطناعي وضمان استخدام هذه التقنية بشكل مسئول ومفيد للمجتمع.

تساؤلات البحث:

يسعى البحث إلى الإجابة على التساؤلات التالية:

١. ما أهم ملامح الواقع الحالي فيما يرتبط بأخلاقيات تصميم وتطوير تطبيقات الذكاء الاصطناعي؟
٢. ما أهم المفاهيم الأخلاقية التي يجب مراعاتها عند تصميم وتطوير واستخدام أنظمة الذكاء الاصطناعي؟
٣. ما أبرز التحديات الأخلاقية التي تنجم عن استخدام الذكاء الاصطناعي في مختلف المجالات؟
٤. ما أهم المعايير الأخلاقية الخاصة بتصميم وتطوير تطبيقات الذكاء الاصطناعي؟

منهج البحث:

وظف البحث الحالي المنهج الوصفي التحليلي، حيث اعتمد على تحليل الدراسات السابقة والتقارير العلمية حول أخلاقيات تصميم وتطوير تطبيقات الذكاء الاصطناعي، فضلا عن تحليل بعض الحالات الواقعية لاستخدام تطبيقات الذكاء الاصطناعي في مختلف السياقات.

مصطلحات البحث:

- **الذكاء الاصطناعي:** يعرف بأنه قدرة الآلة على محاكاة القدرات الذهنية البشرية، مثل التعلم وحل المشكلات واتخاذ القرارات والتكيف والتعامل مع المواقف الجديدة، فضلا عن القيام بالوظائف الأخرى التي تتطلب الذكاء الإنساني (Naqvi, 2020).
- **أخلاقيات الذكاء الاصطناعي:** مجموعة من المبادئ والقيم التي تسعى إلى توجيه تطوير واستخدام تكنولوجيا الذكاء الاصطناعي بشكل مسئول وأخلاقي. وتشمل هذه المبادئ قضايا مثل: الشفافية، والمساءلة، والعدالة، والخصوصية، والأمن، وتهدف إلى ضمان أن الذكاء الاصطناعي يُستخدم بطرق تفيد البشرية وتتجنب إلحاق الضرر (Taddeo & Floridi, 2018).
- **المعايير الأخلاقية لتصميم وتطوير تطبيقات الذكاء الاصطناعي:** هي مجموعة من المبادئ والقيم التي يجب مراعاتها عند تصميم وتطوير هذه التطبيقات لضمان استخدامها بشكل مسئول وأخلاقي. تشمل هذه المعايير قضايا مثل الشفافية، والمساءلة، والعدالة، والخصوصية، والأمن، وتهدف إلى توجيه عملية التطوير لضمان أن الذكاء الاصطناعي يستخدم بطرق تفيد الأفراد وتتجنب إلحاق الضرر بهم (Jobin et al., 2015).

أهم محاور البحث:

يتضمن الجزء التالي مجموعة من المحاور لتوضيح أهمية الاستخدام الأخلاقي لتطبيقات الذكاء الاصطناعي في مختلف مجالات الحياة، فضلا عن توضيح لأهم التحديات وبخاصة التحديات الأخلاقية المرتبطة باستخدام وتوظيف تطبيقات الذكاء الاصطناعي، مع عرض لتجارب الدولة المصرية المرتبطة بصياغة الاستراتيجية الوطنية للذكاء الاصطناعي، مع توضيح لكيفية تعظيم بنود هذه الاستراتيجية في تنفيذ الاستراتيجية الوطنية لمكافحة الفساد، وذلك سعيًا للوصول إلى صياغة قائمة بالمعايير الأخلاقية التي يجب أن تحكم تصميم واستخدام تطبيقات الذكاء الاصطناعي في مختلف المجالات وذلك على النحو التالي:

أولا: أهمية الاستخدام الأخلاقي لتطبيقات الذكاء الاصطناعي:

يشهد مجال أخلاقيات الذكاء الاصطناعي تطورًا ملحوظًا، حيث تتجه الجهود نحو وضع أُطر ومعايير أخلاقية أكثر شمولية ومرونة، ومن أبرز هذه الاتجاهات التركيز على الذكاء الاصطناعي المسئول، وتطوير آليات للمراجعة والتدقيق، والتعاون الدولي، وإشراك الجمهور (Hagendorff, 2020; UNESCO, 2021).

ومن المؤكد أن زيادة استخدام تطبيقات الذكاء الاصطناعي سيؤثر على المهارات المطلوبة في سوق العمل في المستقبل، حيث سيحتاج الأفراد إلى تطوير مهارات جديدة تتناسب مع متطلبات العصر الرقمي، مثل: مهارات التفكير الناقد وحل المشكلات، ومهارات التعاون والتواصل، ومهارات الإبداع والابتكار، ومهارات التكيف والتعلم المستمر.

ويُمكن أن يُساهم استخدام تطبيقات الذكاء الاصطناعي في تحسين مستوى المعيشة للجميع، من خلال زيادة الإنتاجية وخلق فرص عمل جديدة، كما يُمكن أن يسهم في تقليل التفاوت الاجتماعي والاقتصادي، من خلال توفير فرص متساوية للتعليم والتدريب للجميع (Harari, 2016; O'Neil, 2016).

ويمكن للذكاء الاصطناعي أن يؤدي إلى تحسين الكفاءة، مما يؤدي إلى تقليل التكاليف وبالتالي تحقيق فوائد اقتصادية تنعكس على المجتمع، مما يساهم في تحسين حياة الأفراد. وتتوقع المفوضية الأوروبية أن "الذكاء الاصطناعي قد ينتشر عبر عديد من الوظائف والقطاعات الصناعية، مما يعزز الإنتاجية ويحقق نموًا اقتصاديًا قويًا وإيجابيًا. (Craglia, 2018)

تشير أخلاقيات الذكاء الاصطناعي إلى مجموعة من المبادئ والقيم التي توجه تطوير واستخدام الذكاء الاصطناعي بشكل مسئول، وتشمل هذه المبادئ العدالة والشفافية والمسئولية والسلامة (Bryson, Taddeo& Floridi, 2019; Floridi, 2020; Hagendorff, 2020; UNESCO, 2021).

بينما تمثل القوانين المنظمة للذكاء الاصطناعي مجموعة من القواعد واللوائح التي تضعها الحكومات والهيئات التنظيمية لتنظيم استخدام الذكاء الاصطناعي، وتهدف هذه القوانين إلى تحقيق التوازن بين تشجيع الابتكار وحماية المجتمع من المخاطر المحتملة للذكاء الاصطناعي (Goodman & Flaxman, 2017).

ويكمن الفرق الأساسي بين الأخلاقيات والقوانين في طبيعتها الإلزامية، فالأخلاقيات تمثل مبادئ توجيهية غير ملزمة، بينما القوانين قواعد ملزمة يجب على الجميع الالتزام بها. وتعتمد الأخلاقيات على القيم المعنوية التي قد تختلف بين الأفراد والمجتمعات، وهي تُعنى بما "ينبغي" فعله وفقًا لمبادئ الخير والعدالة. في المقابل، القوانين هي قواعد مُلزمة داخل نطاق قضائي معين، تحدد ما "يجب" القيام به، مما يجعلها أكثر صرامةً ووضوحًا في التطبيق. ويمكن التفرقة بين الأخلاقيات والقوانين من وجهة نظر (Flanagan, 2024) فيما يلي:

- تُبنى القرارات الأخلاقية على القناعات الشخصية والأعراف الثقافية، مما يجعلها ذاتية ومتغيرة من سياق لآخر. أما القوانين، فهي موضوعية وتُطبق على جميع الأفراد بغض النظر عن اختلافاتهم الفكرية أو الثقافية، مما يضفي عليها طابعًا موحدًا في التنفيذ.
- الانتهاكات الأخلاقية قد تؤدي إلى عواقب اجتماعية أو مهنية، مثل النبذ المجتمعي أو فقدان السمعة، لكنها لا تحمل بالضرورة آثارًا قانونية. أما مخالفة القوانين، فتترتب عليها عقوبات واضحة، مثل الغرامات المالية أو السجن، ما يجعل أثرها أكثر حدة وإلزامًا.
- تتميز الأخلاقيات بالقدرة على التكيف السريع مع المتغيرات الاجتماعية والتطورات التكنولوجية، ما يجعلها ديناميكية بطبيعتها. في المقابل، فإن القوانين غالبًا ما تكون جامدة، إذ تتطلب عمليات تشريعية معقدة لتحديثها أو تعديلها، مما يؤدي إلى فجوة زمنية بين التغيرات الاجتماعية وسنّ التشريعات المناسبة.
- تلعب الأخلاقيات دورًا إرشاديًا في المجالات التي قد تكون القوانين فيها غير واضحة أو غائبة، لا سيما في القطاعات الناشئة مثل أخلاقيات الذكاء الاصطناعي والهندسة الوراثية. فهي تسد الفراغات القانونية وتوجه السلوك البشري في المواقف غير المنظّمة رسميًا، ما يجعلها عنصرًا تكميليًا للقوانين في صياغة القواعد المجتمعية.

وبالتالي، تشكل الأخلاقيات والقوانين معًا حجر الأساس في الإطار التنظيمي للذكاء الاصطناعي، حيث تعملان على توجيه تطوير واستخدام هذه التقنية بطريقة تضمن تحقيق العدالة، والشفافية، والمسئولية. فالأخلاقيات تحدد المبادئ والقيم التي يجب الالتزام بها عند تصميم وتطبيق أنظمة الذكاء الاصطناعي، مثل: احترام الخصوصية، وتجنب التحيز، وضمان سلامة المستخدمين. أما القوانين، فتعمل على وضع

الأطر القانونية الملزمة التي تنظم استخدام هذه التقنية، من خلال تحديد المسئوليات القانونية، وفرض العقوبات على أي انتهاكات.

وتتضمن أخلاقيات التعامل مع تطبيقات الذكاء الاصطناعي عدة مفاهيم أساسية يمكن توضيحها على النحو التالي (Bostrom, 2014; Zuiderveen et al., 2018; O'Neil, 2016; Taddeo & Floridi, 2018; Zuboff, 2019; Al-Garadi et al.,2020):

- **المسئولية والمساءلة:** حيث تحدد المسئولية من الذي يتحمل نتائج وقرارات أنظمة الذكاء الاصطناعي، سواء كان المطور أو المستخدم أو النظام نفسه، بينما تعني المساءلة وجود آليات لمحاسبة الأطراف المعنية عن تصرفات أنظمة الذكاء الاصطناعي.
- **الشفافية وقابلية التفسير:** حيث يجب أن تكون عمليات اتخاذ القرار في أنظمة الذكاء الاصطناعي واضحة ومفهومة وقابلة للفحص والتفسير.
- **العدالة والإنصاف:** حيث يجب ضمان معاملة جميع الأفراد والمجموعات بإنصاف وعدالة من قبل أنظمة الذكاء الاصطناعي، دون تمييز أو تحيز.
- **الخصوصية وحماية البيانات:** والتي تضمن احترام خصوصية الأفراد وحماية بياناتهم الشخصية من قبل أنظمة الذكاء الاصطناعي.
- **السلامة والأمان:** حيث يجب ضمان سلامة الأفراد والمجتمع من أي ضرر قد ينجم عن استخدام أنظمة الذكاء الاصطناعي، مع حماية هذه الأنظمة من الهجمات السيبرانية والتلاعب، ويوضح الشكل التالي أخلاقيات التعامل مع تطبيقات الذكاء الاصطناعي.

شكل (٢) أخلاقيات التعامل مع تطبيقات الذكاء الاصطناعي (Taddeo & Floridi, 2018)

وهناك عديد من السبل التي يمكن من خلالها تعزيز التعامل الأخلاقي مع تطبيقات الذكاء الاصطناعي، من بينها: وضع معايير أخلاقية واضحة، وتضمين الأخلاقيات في تصميم وتطوير الأنظمة، وتدريب المتخصصين، وتوعية المستخدمين، وتشجيع البحث العلمي. وتعد مراعاة البعد الأخلاقي في

تصميم وتطوير أنظمة الذكاء الاصطناعي أمرًا ضروريًا لضمان استخدام هذه التقنية بشكل مسئول ومفيد للمجتمع. ويلخص الشكل التالي أهم السبل التي تعزز التعامل الأخلاقي مع تطبيقات الذكاء الاصطناعي:

شكل (٣) سبل تعزيز التعامل الأخلاقي مع تطبيقات الذكاء الاصطناعي (Al-Garadi et al., 2020)

ولذلك تُعد أخلاقيات التعامل مع تطبيقات الذكاء الاصطناعي عنصرًا حاسمًا لضمان تطوير وتطبيق هذه التقنيات بطرق تعزز العدالة، وتضمن السلامة، وتفيد المجتمع ككل. ويمكن تعزيز التعامل الأخلاقي مع الذكاء الاصطناعي من خلال نشر الوعي بالمعايير الأخلاقية المختلفة، مع تحميل المستخدمين المسئولية الكاملة وراء اقتباس المعلومات من أدوات الذكاء الاصطناعي وضرورة التحقق من مصداقيتها (McKendrick, 2022).

ثانيا: أهم التحديات الأخلاقية المرتبطة باستخدام تطبيقات الذكاء الاصطناعي:

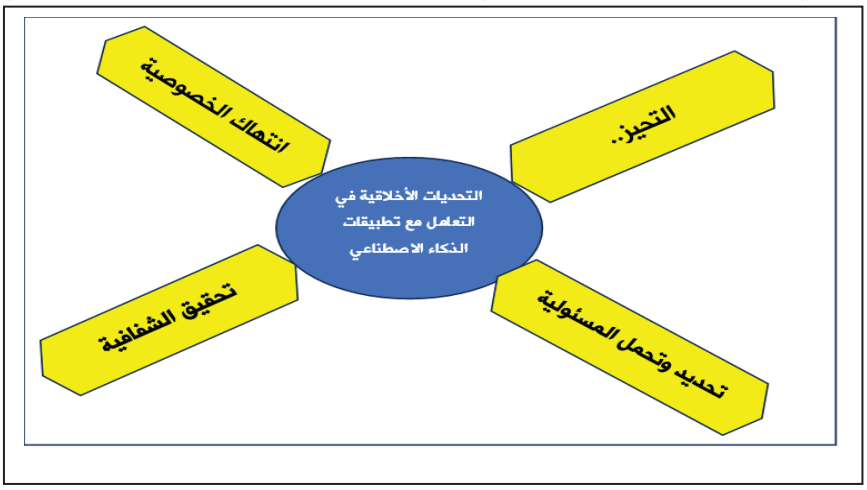
يُمكن أن يلعب الذكاء الاصطناعي دورًا مُزدوجًا في انتشار الأخبار الزائفة والمعلومات المضللة، فمن ناحية؛ يُمكن أن يُستخدم الذكاء الاصطناعي لإنتاج ونشر الأخبار الزائفة بسرعة وفعالية أكبر، مما يزيد من صعوبة اكتشافها والتحقق منها، ومن ناحية أخرى؛ يُمكن أن يُستخدم الذكاء الاصطناعي أيضًا لتطوير أدوات وتقنيات تُساعد على كشف الأخبار الزائفة ومكافحتها (Marwick& Lewis, 2017; Wardle & Derakhshan, 2017).

كما أن تقنيات التزييف من أبرز التحديات الأخلاقية التي تواجه الإعلام في عصر الذكاء الاصطناعي، فباستخدام هذه التقنيات، يُمكن إنشاء مقاطع فيديو أو صور مزيفة تبدو واقعية تمامًا، مما قد يُستخدم للتضليل أو تشويه سمعة الأفراد والمؤسسات، وتُثير تقنيات التزييف العميق عديدًا من المخاوف الأخلاقية، مثل: انتهاك الخصوصية والتضليل الإعلامي وتشويه السمعة. ويواجه استخدام تطبيقات الذكاء الاصطناعي عديدًا من التحديات الأخلاقية، من بينها ما يلي:

(Mittelstadt et al., 2016 ; Obermeyer et al., 2019 ; Rahwan et al., 2019 ; Brundage et al., 2020) :

- **التحيز**: حيث قد تتضمن أنظمة الذكاء الاصطناعي تحيزات موجودة في البيانات التي تم تضمينها بالنظام، مما يؤدي إلى نتائج غير عادلة.
- **تحديد وتحمل المسئولية**: حيث يثار تساؤل حول من يتحمل مسئولية الأضرار الناجمة عن أنظمة الذكاء الاصطناعي.
- **تحقيق الشفافية**: حيث قد تكون عمليات اتخاذ القرار في بعض أنظمة الذكاء الاصطناعي معقدة للغاية، مما يجعل من الصعب فهمها وشرحها.
- **انتهاك الخصوصية**: حيث يثير جمع وتحليل كميات هائلة من البيانات الشخصية مخاوف بشأن الخصوصية.

ويلخص الشكل التالي أهم التحديات الأخلاقية في التعامل مع تطبيقات الذكاء الاصطناعي:



شكل (٤) التحديات الأخلاقية في التعامل مع تطبيقات الذكاء الاصطناعي (Brundage et al., 2020)

كما أن المسئولية عن الأضرار الناجمة عن استخدام تطبيقات الذكاء الاصطناعي من القضايا المعقدة، حيث يصعب تحديد المسئول بشكل قاطع في عديد من الحالات، فالذكاء الاصطناعي ليس شخصًا طبيعيًا أو اعتباريًا، بل هو نظام تقني يعتمد على الخوارزميات والبيانات (Bostrom, 2014).

وهناك عدة أطراف يُمكن أن تتحمل المسئولية عن أخطاء الذكاء الاصطناعي، منها المطورون والمصنعون والمشغلون والمالكون. ونظرًا لتعقيد المسئولية عن أخطاء الذكاء الاصطناعي، تبرز الحاجة إلى وضع تشريعات تنظم استخدام هذه التقنية بشكل واضح وفعال، حيث يجب أن تُحدد هذه التشريعات المسئوليات بوضوح، وأن تتضمن آلياتٍ للتعويض عن الأضرار الناجمة عن أخطاء الذكاء الاصطناعي، بالإضافة إلى ذلك، يجب أن تتناول التشريعات قضايا أخرى مثل: حماية البيانات والشفافية والمساءلة. ولقد بدأت بعض الدول في وضع نماذج للمساءلة الأخلاقية والقانونية في مجال الذكاء الاصطناعي، من بين هذه الدول الاتحاد الأوروبي والولايات المتحدة وسنغافورة.

وتحتاج آليات اتخاذ القرار في أنظمة الذكاء الاصطناعي إلى قدر كبير من الشفافية، حيث يجب أن يكون المستخدمون والمطورون على حد سواء قادرين على فهم كيفية عمل هذه الأنظمة، وكيف تصل إلى النتائج التي تقدمها وتكمن أهمية الشفافية في أنها تساعد على بناء الثقة، وتحديد الأخطاء والتحيزات، وتحسين الأداء (Doshi-Velez & Kim, 2017).

ويمكن أن يؤدي التحيز في تصميم تطبيقات الذكاء الاصطناعي إلى عواقب وخيمة تؤثر على الأفراد والمجتمعات بطرق مختلفة. فعند برمجة الأنظمة الذكية بالاعتماد على بيانات غير متوازنة أو متحيزة، قد تنعكس هذه التحيزات في القرارات التي تتخذها المؤسسة، مما يؤدي إلى التمييز ضد فئات معينة مثل: الأقليات العرقية أو الاجتماعية. كما يمكن أن تتفاقم أزمة التمييز وعدم المساواة في مجالات مثل: التوظيف، والرعاية الصحية، والعدالة الجنائية، حيث قد تتخذ الأنظمة قرارات غير عادلة بناءً على أنماط غير موضوعية في البيانات. لذلك، يعد التصدي لهذه المشكلة من خلال تطوير تطبيقات ذكية أكثر عدالة، وتحسين جودة البيانات أمرًا ضروريًا لضمان استخدام الذكاء الاصطناعي بشكل مسئول ومنصف.

وتُعد العدالة في الوصول إلى تطبيقات الذكاء الاصطناعي واستخدامها من القضايا المهمة الأخرى، حيث يجب أن يكون الجميع قادرين على الاستفادة من هذه التقنية، بغض النظر عن خلفياتهم أو قدراتهم، وتشمل العدالة في هذا السياق توفير الوصول الشامل وضمان الاستخدام العادل (Goodman & Flaxman, 2017).

كما يُعد تحقيق التوازن بين الأمن والحرية الفردية من التحديات الكبرى التي تواجه المُجتمعات في العصر الرقمي، فمن ناحية، يجب علينا حماية أنفسنا ومجتمعاتنا من المخاطر الأمنية، ومن ناحية أخرى، يجب علينا احترام حقوق الأفراد وحرياتهم. ولتحقيق هذا التوازن، يجب علينا وضع قوانين ولوائح واضحة، كالشفافية والمساءلة، والرقابة والتطوير.

كما أن أخلاقيات جمع البيانات واستخدامها في الذكاء الاصطناعي من القضايا المهمة التي يجب مراعاتها، حيث يجب أن يتم جمع البيانات بطريقة شفافة ومسئولة، كما يجب أن يتم استخدامها فقط للأغراض التي جُمعت من أجلها. وتجدر الإشارة إلى أن عدم مراعاة أخلاقيات جمع البيانات واستخدامها قد يسفر عنه عديد من المخاطر المرتبطة بانتهاك خصوصية الأفراد، والتي قد تؤثر على حياتهم بشكل مباشر، فعلى سبيل المثال: المراقبة المستمرة للأفراد بدون إذن مسبق قد تؤدي إلى تقييد حريتهم الشخصية والتأثير على سلوكهم بسبب الشعور الدائم بأنهم تحت المراقبة. فضلا عن مخاطر أخرى مثل: سرقة الهوية الشخصية للأفراد أو بياناتهم الخاصة؛ مما قد ينتج عنه استخدام غير قانوني للمعلومات في الاحتيال المالي (Zuboff, 2019).

ويُثير موضوع تأثير استخدام تطبيقات الذكاء الاصطناعي في العمل جدلًا واسعًا، حيث يرى البعض أنه سيؤدي إلى فقدان وظائف كثيرة، بينما يرى البعض الآخر أنه سيخلق فرص عمل جديدة، ومن المتوقع

أن يؤدي الذكاء الاصطناعي إلى رقمنة بعض الوظائف، خاصة تلك التي تتطلب مهام روتينية ومتكررة، كما سيخلق الذكاء الاصطناعي أيضًا فرص عمل جديدة، خاصة في مجالات تطوير وتصميم وصيانة أنظمة الذكاء الاصطناعي، وتحليل البيانات، وغيرها من المجالات التي تعتمد على التكنولوجيا المتقدمة (Agrawal et al., 2018; Brynjolfsson & McAfee, 2014; Ford, 2015)

ثالثا: الاستراتيجية الوطنية المصرية للذكاء الاصطناعي:

إن اتخاذ القرار الأخلاقي باستخدام تطبيقات الذكاء الاصطناعي من الأمور التي تثير جدلًا واسعًا في الأوساط العلمية والفكرية، حيث يرى البعض أن الذكاء الاصطناعي، بحكم طبيعته التقنية، لا يُمكن أن يتجاوز حدود المنطق والبيانات التي تعتمد عليها عمليات التحليل Data analysis، ويفتقر إلى القدرة على فهم القيم الأخلاقية المعقدة والمتغيرة، بينما يرى البعض الآخر أن الذكاء الاصطناعي، مع تطور قدراته الحسابية والتحليلية، قد يكون قادرًا على اتخاذ قرارات أخلاقية أفضل من البشر في بعض الحالات، خاصة عندما يتعلق الأمر بتقييم كميات هائلة من البيانات واتخاذ قرارات سريعة. وتمثل برمجة القيم الإنسانية أحد التحديات الكبرى التي تواجه تطبيقات الذكاء الاصطناعي، فالقيم الأخلاقية غالبًا ما تكون ضمنية ومتغيرة، وتتأثر بعديد من العوامل الثقافية والاجتماعية والشخصية، وتحويل هذه القيم إلى قواعد رقمية قابلة للبرمجة ليس بالأمر السهل، وقد يؤدي إلى تبسيط مُخل أو تحريف للقيم الإنسانية الحقيقية.

كما تزداد أهمية تقصي إمكانية اتخاذ القرارات الأخلاقية عن طريق تطبيقات الذكاء الاصطناعي في المجالات الحساسة مثل: الطب والقضاء والأمن، حيث قد يكون للقرارات المُتخذة بواسطة الذكاء الاصطناعي تأثيرات عميقة على حياة الأفراد والمجتمع ككل، ففي المجال الطبي، على سبيل المثال، قد يحتاج الذكاء الاصطناعي إلى اتخاذ قرارات صعبة بشأن توزيع الموارد الطبية المحدودة، أو تحديد أفضل العلاجات لحالات مُعقدة، وفي المجال القضائي، قد يُستخدم الذكاء الاصطناعي في تحليل الأدلة وتقديم التوصيات للقضاة، وفي المجال الأمني، قد يُستخدم الذكاء الاصطناعي في مُراقبة الحدود وتحديد التهديدات المُحتملة.

ولذا فقد أطلقت مصر الاستراتيجية الوطنية للذكاء الاصطناعي بهدف الاستفادة من هذه التكنولوجيا المتقدمة في تحقيق التنمية المستدامة في مختلف القطاعات. وتركز الاستراتيجية على تطوير القدرات المحلية في مجال الذكاء الاصطناعي، وتشجيع البحث والابتكار، وبناء البنية التحتية اللازمة لدعم تطبيقات الذكاء الاصطناعي. كما تهدف إلى تعزيز التعاون الدولي في هذا المجال وتبادل الخبرات مع الدول المتقدمة (وزارة الاتصالات وتكنولوجيا المعلومات، ٢٠٢١).

وتتضمن الاستراتيجية الوطنية للذكاء الاصطناعي في مصر عدة محاور رئيسية، منها تطوير المهارات والكفاءات في مجال الذكاء الاصطناعي، وتشجيع الاستثمار في الشركات الناشئة التي تعمل في هذا المجال، وتطوير البنية التحتية الرقمية اللازمة لدعم تطبيقات الذكاء الاصطناعي. كما تركز

الاستراتيجية على تعزيز الثقة في استخدامات الذكاء الاصطناعي وضمان الشفافية والمساءلة في تطوير وتطبيق هذه التكنولوجيا (الجهاز المركزي للتعبئة العامة والإحصاء، ٢٠٢٢).

كما تهدف الاستراتيجية الوطنية للذكاء الاصطناعي في مصر إلى تحقيق عدة أهداف استراتيجية، منها تحسين جودة الخدمات الحكومية، وتعزيز الابتكار في القطاعات الاقتصادية المختلفة، وتحسين مستوى التعليم والصحة، وتعزيز الأمن السيبراني. وتسعى مصر من خلال هذه الاستراتيجية إلى أن تصبح مركزاً إقليمياً للذكاء الاصطناعي، وأن تساهم في تطوير هذه التكنولوجيا على المستوى العالمي (المجلس الوطني للذكاء الاصطناعي، ٢٠٢٣).

رابعا: آليات توظيف تطبيقات الذكاء الاصطناعي في مكافحة الفساد:

أطلقت مصر الاستراتيجية الوطنية لمكافحة الفساد بهدف تعزيز النزاهة والشفافية في مختلف مؤسسات الدولة، ومكافحة جميع أشكال الفساد المالي والإداري. تركز الاستراتيجية على تطوير الأطر التشريعية والمؤسسية لمكافحة الفساد، وتفعيل دور الأجهزة الرقابية، وتعزيز التعاون الدولي في هذا المجال. كما تهدف إلى نشر الوعي المجتمعي بأهمية مكافحة الفساد وتفعيل دور المجتمع المدني في هذا الإطار (هيئة الرقابة الإدارية، ٢٠٢٣).

وتتضمن الاستراتيجية الوطنية لمكافحة الفساد في مصر عدة محاور رئيسية، من بينها: تعزيز الشفافية وتطوير آليات المراجعة والتدقيق المالي، وتفعيل دور الأجهزة الرقابية في الكشف عن جرائم الفساد وملاحقة مرتكبيها. كما تركز الاستراتيجية على تطوير منظومة الخدمة المدنية وتحسين إجراءات التعيين والترقية، وتفعيل المساءلة عن المخالفات والتجاوزات. كما تهدف الاستراتيجية الوطنية لمكافحة الفساد في مصر إلى تحقيق عدة أهداف استراتيجية، من بينها: الحد من انتشار الفساد في مختلف القطاعات، وتعزيز الثقة في مؤسسات الدولة، وتحسين مناخ الاستثمار وجذب الاستثمارات الأجنبية، وتحقيق التنمية المستدامة. وتسعى مصر من خلال هذه الاستراتيجية إلى أن تصبح دولة رائدة في مجال مكافحة الفساد، فضلا عن المساهمة في الجهود الدولية المبذولة لمكافحة هذه الظاهرة.

كما يلعب الذكاء الاصطناعي دوراً متزايد الأهمية في مكافحة الفساد في مصر، حيث يمكن لأدواته أن تساعد في تحليل كميات هائلة من البيانات للكشف عن الأنماط المشبوهة والتنبؤ بوجود مخالفات محتملة. على سبيل المثال: يمكن استخدام تقنيات تعلم الآلة لتحليل البيانات المالية والإدارية للكشف عن عمليات الفساد والتلاعب بالمال العام. كما يمكن استخدام الذكاء الاصطناعي في تحليل العقود الحكومية ومراقبة المشاريع للتأكد من نزاهة العمليات وعدم وجود تلاعب أو فساد (وزارة الاتصالات وتكنولوجيا المعلومات، ٢٠٢٣).

وتتنوع أدوات الذكاء الاصطناعي التي يمكن أن تستخدم في مكافحة الفساد، فمنها ما يعتمد على تحليل البيانات الضخمة للكشف عن المخالفات، ومنها ما يستخدم تقنيات معالجة اللغة الطبيعية لتحليل النصوص والكشف عن الرسائل المشبوهة. كما توجد أدوات تعتمد على الرؤية الحاسوبية لتحليل الصور ومقاطع الفيديو للكشف عن عمليات التزوير والتلاعب. فضلا عن ذلك، يمكن استخدام الذكاء الاصطناعي في تطوير أنظمة ذكية للإنذار المبكر للكشف عن المخاطر المحتملة للفساد واتخاذ الإجراءات الوقائية اللازمة. ومن ذلك وعلى سبيل المثال، تمكنت هيئة الرقابة الإدارية باستخدام تحليلات الذكاء الاصطناعي لقواعد بيانات هيئة الرقابة الادارية من ضبط تشكيل عصابي تخصص في اصطناع وتزوير مستندات حكومية تفيد أحقية بعض المواطنين في الاستفادة من امتيازات برامج الرعاية الاجتماعية المقررة للأشخاص ذوي الإعاقة بالمخالفة للحقيقة، وتمكينهم من الحصول على معاشات استثنائية واستيراد السيارات دون سداد الرسوم الجمركية (هيئة الرقابة الإدارية، ٢٠٢٣).

وتساهم أدوات الذكاء الاصطناعي في مكافحة الفساد في مصر من خلال عدة جوانب، منها تحسين كفاءة الأجهزة الرقابية في الكشف عن المخالفات، وتقليل احتمالية وقوع الفساد من خلال التنبؤ بالمخاطر المحتملة. كما تساعد هذه الأدوات في تعزيز المساءلة ومكافحة الإفلات من العقاب، حيث يمكن استخدامها في تتبع الأموال العامة واستعادة الأموال المنهوبة. وتسعى مصر إلى تعزيز استخدام الذكاء الاصطناعي في مكافحة الفساد من خلال تطوير القدرات المحلية في هذا المجال، وتشجيع التعاون بين القطاعين العام والخاص في تطوير وتطبيق هذه الأدوات (الجهاز المركزي للمحاسبات، ٢٠٢١).

خامسا: الأساس النظري للبحث:

إحدى النظريات التي يمكن أن تشكل أساسًا نظريًا لهذا البحث هي نظرية الحوكمة الأخلاقية للذكاء الاصطناعيEthical AI Governance Theory تفترض هذه النظرية أن تطوير وتطبيق أنظمة الذكاء الاصطناعي يجب أن يخضع لإطار حوكمة شامل يدمج مبادئ أخلاقية مثل العدالة، الشفافية، والمساءلة. وفقًا لهذه النظرية، فإن المؤسسات والجهات التنظيمية يجب أن تضع سياسات صارمة لضمان توافق الأنظمة الذكية مع القيم الإنسانية والمجتمعية .(Floridi et al., 2018)

كما تؤكد على ضرورة تطوير أنظمة ذكاء اصطناعي قابلة للتفسير وتجنب التحيزات التي قد تؤثر على مخرجاته (Jobin, Ienca, & Vayena, 2019).

كما تدعم نظرية الحوكمة الأخلاقية مبدأ "التصميم الأخلاقي المسبقEthics by Design، حيث يُدمج التفكير الأخلاقي في جميع مراحل تطوير أنظمة الذكاء الاصطناعي، بدءًا من جمع البيانات وحتى مرحلة التطبيق الفعلي. (Mittelstadt, 2019)

وتؤكد الأبحاث أن تطبيق هذه النظرية يؤدي إلى تعزيز الثقة العامة في تقنيات الذكاء الاصطناعي ويقلل من المخاطر الاجتماعية المرتبطة باستخدامها، مثل انتهاك الخصوصية أو اتخاذ قرارات غير عادلة ضد مجموعات معينة (Morley et al., 2020).

بالإضافة إلى ذلك، ترتبط هذه النظرية بالمبادئ التوجيهية التي أصدرتها جهات دولية مثل اليونسكو والمفوضية الأوروبية، والتي تهدف إلى وضع معايير قانونية وأخلاقية تضمن استخدام الذكاء الاصطناعي بطرق تحترم حقوق الإنسان والقيم المجتمعية. وبهذا، تعد نظرية الحوكمة الأخلاقية للذكاء الاصطناعي إطارًا نظريًا مناسبًا لفهم المعايير الأخلاقية في تصميم وتطوير تطبيقات الذكاء الاصطناعي في العصر الرقمي(European Commission, 2020) .

سادسا: تحليل الموقف الحالي فيما يتعلق بأخلاقيات نظم الذكاء الاصطناعي:

تُعد الشفافية من أهم المبادئ الأخلاقية في تطوير وتطبيقات الذكاء الاصطناعي، إلا أن تطبيقها في مصر يواجه تحديات تتعلق بقلة الإفصاح عن خوارزميات الذكاء الاصطناعي المستخدمة في القطاعات المختلفة، مما قد يؤدي إلى غياب المساءلة والتأثير السلبي على اتخاذ القرار. وفقًا لزوبوف Zuboff (2019) وبري ودينوف Brey & Dainow (2023) فإن تعزيز الشفافية يتطلب سياسات واضحة تُلزم مطوري التطبيقات بالإفصاح عن كيفية عمل الأنظمة الذكية وتأثيرها المحتمل على المستخدمين، وهو ما لا يزال في طور التطوير في البيئة المصرية الرقمية.

من ناحية العدالة، فإن هناك قلقًا متزايدًا بشأن تحيزات الذكاء الاصطناعي التي قد تؤثر على القرارات في مجالات مثل التوظيف والتعليم والخدمات الحكومية. تشير دراسات "أندرسون" Anderson (2007) و"بوستروم" Bostrom (2014) إلى أن هذه التحيزات قد تنجم عن عدم تمثيل البيانات بشكل عادل لجميع الفئات الاجتماعية، مما يؤدي إلى مخرجات غير متوازنة. في مصر، لا تزال هناك جهود محدودة لضمان عدالة البيانات المستخدمة في تدريب النماذج الذكية، ما يستدعي تدخلات قانونية وتقنية لضبط معايير تطوير التطبيقات وضمان استخدامها بإنصاف.

أما فيما يخص الخصوصية والأمن السيبراني، فإن الاعتماد المتزايد على تطبيقات الذكاء الاصطناعي في مصر يثير مخاوف تتعلق بحماية البيانات الشخصية وضمان استخدامها بشكل مسئول. وفقًا لـ "كاث"Cath (2018) وبري و"دينوف" Brey & Dainow (2023) فإن تعزيز أمن البيانات يتطلب وجود قوانين واضحة وإجراءات فعالة لحماية المعلومات من الاختراق أو سوء الاستخدام. في مصر، تم سن قوانين لحماية البيانات الشخصية، إلا أن التطبيق الفعلي لهذه القوانين يواجه تحديات تتعلق بالقدرة على تنفيذها بفعالية، خاصة مع تنامي التقنيات التي تعتمد على تحليل البيانات الضخمة.

سابعا: اشتقاق قائمة المعايير الأخلاقية لتصميم وتطوير تطبيقات الذكاء الاصطناعي:

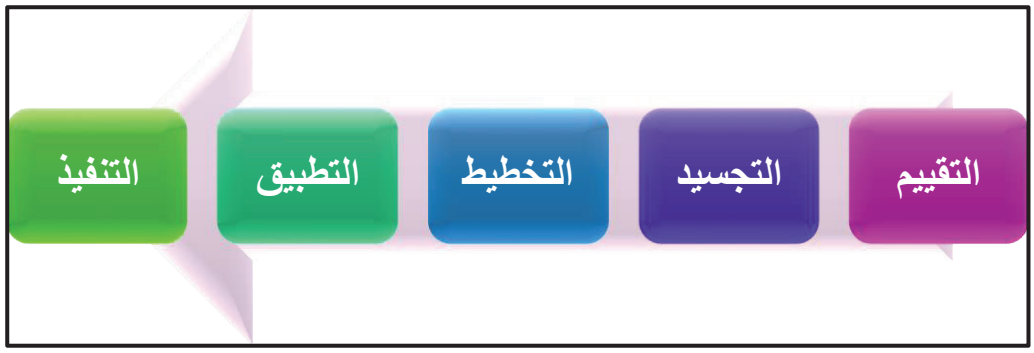
تُعد المعايير الأخلاقية لتصميم وتطوير وإنتاج تطبيقات الذكاء الاصطناعي ذات أهمية قصوى في العصر الحديث، حيث أصبحت هذه التقنية جزءاً لا يتجزأ من حياتنا. يجب أن يتم تصميم هذه التطبيقات بطريقة تضمن احترام حقوق الإنسان والحريات الأساسية، وعدم التمييز أو التحيز ضد أي فئة من الناس. كما يجب أن تكون التطبيقات شفافة وقابلة للمساءلة، وأن يكون المستخدمون على دراية بكيفية عملها واتخاذ القرارات (Anderson & Anderson, 2007).

تشمل المعايير الأخلاقية لتصميم تطبيقات الذكاء الاصطناعي مجموعة واسعة من القضايا، مثل: الخصوصية والأمن والمسئولية والشفافية. ويجب أن يتم تصميم التطبيقات بطريقة تحمي خصوصية المستخدمين وبياناتهم الشخصية، كما يجب أن يكون هناك تحديد واضح للمسئولية في حالة وقوع أضرار نتيجة لاستخدام تلك التطبيقات.(Bostrom, 2014)

كما يجب أن يتم تصميم تطبيقات الذكاء الاصطناعي بطريقة تضمن عدم التمييز أو التحيز ضد أي فئة من الناس. ويجب أن تكون التطبيقات عادلة ومنصفة، وألا تتسبب في أي نوع من أنواع التمييز أو الظلم. كما يجب أن يتم اختبار التطبيقات بعناية للتأكد من أنها لا تحتوي على أي تحيزات أو أخطاء قد تؤدي إلى التمييز .(O'Neil, 2016)

كما أن المسئولية الاجتماعية من أهم المعايير الأخلاقية لتصميم تطبيقات الذكاء الاصطناعي. ويجب أن يكون المصممون على دراية بالآثار الاجتماعية المحتملة لتطبيقاتهم، وأن يكونوا على استعداد لتحمل المسئولية عن أي أضرار قد تنجم عن استخدام تطبيقاتهم.(Russell & Norvig, 2021)

وفيما يلي عرض لخطوات تطبيق الأخلاقيات في تصميم وإنتاج تطبيقات الذكاء الاصطناعي (Brey& Dainow, 2023):



شكل (٥) خطوات تطبيق الأخلاقيات في تصميم وإنتاج تطبيقات الذكاء الاصطناعي (Brey& Dainow, 2023)

- **التقييم**: أولاً، يتم تقييم أهداف النظام مقابل القيم الأخلاقية الأساسية. حتى لا يتم انتهاك أي من القيم الأخلاقية الأساسية.

- **التجسيد:** يتم تحويل القيم الأخلاقية إلى مؤشرات محددة يجب أن تتوافر في نظام الذكاء الاصطناعي المستخدم.
- **التخطيط:** يشمل التخطيط عملية تحويل متطلبات التصميم الأخلاقي العامة إلى إجراءات محددة يتم تنفيذها أثناء عملية التصميم. ويتم تنفيذ هذه المتطلبات من خلال وسائل مختلفة، بما في ذلك وظائف النظام، وهياكل البيانات، والتدابير التنظيمية.
- **التطبيق:** يشمل تحديد الطريقة التي سيتم من خلالها معالجة كل متطلب أخلاقي. ويساعد ذلك على توفير نموذج عام للتطوير، مما يسهل دمج المتطلبات الأخلاقية في مراحل التصميم والتنفيذ المختلفة، وفي بعض الحالات يتم إعادة عملية التطبيق.
- **التنفيذ:** يمثل الخطوة الأخيرة في تطبيق الأخلاقيات حسب التصميم أثناء عملية التطوير. وعلى الرغم من أن هذه الخطوة يمكن وصفها ببساطة، إلا أنها تشكل الجزء الأكبر من العمل عند تطوير نظام ذكاء اصطناعي.

في بعض الحالات، قد تتطلب المتطلبات الأخلاقية إضافة وظائف جديدة، بينما في حالات أخرى، قد تعمل كقيود تحدد الوظائف المسموح بها. كما أن العديد من هذه المتطلبات تستدعي عمليات تنظيمية إضافية. على سبيل المثال، لتجنب التحيز في البرمجة، يجب إجراء تقييم رسمي للتحيز في البيانات قبل استخدامها.

خطوات إعداد قائمة بالمعايير الأخلاقية لتصميم وتطوير تطبيقات الذكاء الاصطناعي:

تم إعداد قائمة المعايير الأخلاقية لتصميم وتطوير تطبيقات الذكاء الاصطناعي من خلال الخطوات التالية:

أ. **تحديد الهدف من بناء قائمة المعايير:**

التحديد الدقيق لقائمة المعايير الأخلاقية لتصميم وتطوير تطبيقات الذكاء الاصطناعي ليستفيد منها المبرمجون والمطورون والتي تنعكس بشكل إيجابي أيضا على جميع المستخدمين.

ب. **اشتقاق المعايير ومؤشراتها:**

تم الرجوع إلى عدد من الأدبيات والدراسات مثل: (Jobin et al, 2019; Brey& Dainow,2023) لإعداد الصورة المبدئية لقائمة المعايير.

ج. **التحقق من صدق قائمة المعايير:**

بعد إعداد القائمة في صورتها المبدئية (ملحق:١)، تم عرض الاستبيان بما تضمنه من معايير على مجموعة من السادة المحكمين (ملحق:٢) والمتخصصين في مجال الحاسبات والمعلومات، لإبداء آرائهم حول:

- مدى أهمية المجالات الرئيسة لقائمة المعايير.
- مدى مناسبة المعايير لتحقيق البيئة لأهدافها.
- مدى صحة الصياغة اللغوية لبنود قائمة المعايير.
- التعديل في القائمة بالإضافة أو الحذف.

د. التحقق من ثبات قائمة المعايير:

تم تعديل قائمة المعايير في ضوء آراء السادة المحكمين، وتم حساب نسبة اتفاق استجابة السادة المحكمين على بنود الاستبيان من خلال معادلة كوبر التالية:

نسبة الاتفاق= عدد مرات الاتفاق/ (عدد مرات الاتفاق+ عدد مرات الاختلاف) X١٠٠

وتم الإبقاء على المعايير التي وصلت نسبة الاتفاق فيها إلى ٨٥٪ واستبعاد المعايير التي قلت عن هذه النسبة، كما اتفق أغلب المحكمين على تعديل الصياغة اللغوية لبعض المعايير، وحذف بعضها، وإعادة صياغة بعض المؤشرات (ملحق:٣).

هـ. الصورة النهائية لقائمة المعايير:

تم إجراء كافة التعديلات المطلوبة وبذلك أصبحت القائمة في صورتها النهائية مكونة من (١٥) معيارًا و(٧٣) مؤشرًا (ملحق:٤)

نتائج البحث:

استهدف البحث التحقق من مجموعة من الفروض المتعلقة بالمعايير الأخلاقية لتطوير تطبيقات الذكاء الاصطناعي، وجاءت النتائج كما يلي:

نتائج اختبار الفرض الأول للبحث:

ينص الفرض الأول للبحث على: وجود معايير أخلاقية رئيسية يجب مراعاتها عند تطوير تطبيقات الذكاء الاصطناعي ولقد أكدت نتائج مسح الدراسات السابقة صحة هذا الفرض، حيث تم التوصل إلى قائمة مكونة من (١٥) معيارًا أخلاقيًا رئيسيًا، تشمل المسئولية، المساءلة، الشفافية، العدالة، الخصوصية، والاستدامة البيئية، كما أظهرت النتائج أن الالتزام بهذه المعايير يساعد في تحسين موثوقية التطبيقات وكفاءتها، إذ إن الالتزام بالمسئولية والمساءلة والشفافية يعزز من ثقة المستخدمين، بينما يساهم التركيز على العدالة والخصوصية في تقليل المشكلات الأخلاقية والقانونية.

نتائج اختبار الفرض الثاني للبحث:

أشار الفرض الثاني للبحث إلى الحاجة إلى مزيد من الأبحاث لفهم انعكاسات هذه المعايير على المطورين والمستخدمين، وأظهر التحليل النظري لنتائج البحوث والدراسات ذات الصلة صحة هذا الفرض

مع التأكيد على ضرورة تعميق الدراسة حول تأثير مراعاة المعايير الأخلاقية على تجربة المستخدمين، وكيفية استجابة المطورين لمتطلبات الامتثال لهذه المعايير.

وللإجابة عن أسئلة البحث:

- للإجابة عن السؤال الأول للبحث والخاص بتحديد أهم المعايير الأخلاقية التي يجب مراعاتها عند تطوير تطبيقات الذكاء الاصطناعي. توصل البحث إلى قائمة مكونة من (١٥) معيارًا، من أبرزها: (المسئولية، والمساءلة، والشفافية، والعدالة، والخصوصية، الاستدامة البيئية).
- ولتحديد كيف يمكن ضمان التزام تطبيقات الذكاء الاصطناعي بهذه المعايير، أوضحت الدراسة التحليلية للدراسات السابقة أنه يجب أن يتم ذلك من خلال وضع سياسات واضحة تحكم عملية التطوير، وإجراء اختبارات دورية للتحقق من التزام التطبيقات بالمعايير الأخلاقية، إضافةً إلى إشراك المستخدمين والخبراء في تقييم هذه التطبيقات.
- وفيما يرتبط بتحديد الآثار المترتبة على الالتزام أو عدم الالتزام بهذه المعايير، أظهرت النتائج أن الالتزام بالمعايير يؤدي إلى تحسين موثوقية التطبيقات وزيادة ثقة المستخدمين، بينما قد يؤدي تجاهلها إلى مشكلات قانونية، وأخلاقية، وحتى تجارية، مثل فقدان ثقة العملاء وظهور تحيزات في قرارات الذكاء الاصطناعي.

وتعكس نتائج البحث التوجه العالمي نحو تطوير ذكاء اصطناعي مسئول، حيث تسعى الحكومات والمؤسسات الأكاديمية والتكنولوجية إلى وضع معايير أخلاقية لتنظيم هذا المجال. ويعود التركيز على الشفافية والعدالة إلى المخاوف المتزايدة من التحيز والتأثير غير العادل لبعض التطبيقات، كما أن إدراج الاستدامة البيئية يعكس تزايد الوعي بأثر التكنولوجيا على البيئة. وأشار البحث إلى أن الالتزام بهذه المعايير لا يجب أن يكون مجرد إجراء نظري، بل يجب أن يترجم إلى سياسات واضحة وإجراءات تقنية خلال مراحل التصميم والتطوير، وهذا يتطلب تعاونًا بين المطورين، المستخدمين، وصانعي السياسات لضمان تحقيق توازن بين الابتكار والمسئولية.

أهم **التوصيات والاستخلاصات:**

في ضوء ما تم مسحه من دراسات سابقة وما تم التوصل إليه من اشتقاق قائمة بالمعايير الأخلاقية لتصميم واستخدام تطبيقات الذكاء الاصطناعي يمكن التوصية بما يلي:

- يتعين على المنظمات الدولية والحكومات العمل على وضع معايير أخلاقية واضحة وشاملة تنظم عملية تصميم وتطوير واستخدام أنظمة الذكاء الاصطناعي، بحيث تضمن تحقيق العدالة والنزاهة في هذه التقنيات. ومن الضروري أن تتسم أنظمة الذكاء الاصطناعي بالشفافية والقابلية للتفسير، مما يسمح بفهم آليات عملها واتخاذ قرارات مستنيرة بشأن استخدامها، إلى جانب وضع آليات واضحة للمساءلة عن أي أضرار أو تداعيات سلبية قد تترتب عليها.

- يجب اتخاذ إجراءات صارمة لحماية البيانات الشخصية، ومنع أي استغلال أو إساءة استخدام لها في تطبيقات الذكاء الاصطناعي، وذلك من خلال فرض سياسات خصوصية قوية وآليات رقابية فعالة. ولتحقيق الاستخدام الأخلاقي والمسئول لهذه التقنيات لابد من تعزيز التعاون الدولي بين الحكومات والمنظمات الدولية والمجتمع المدني، بهدف وضع سياسات موحدة تسهم في توجيه الذكاء الاصطناعي نحو المصلحة العامة.
- ينبغي إشراك الجمهور في المناقشات المتعلقة بأخلاقيات الذكاء الاصطناعي، من خلال تنظيم حوارات مفتوحة وفعاليات توعوية، وذلك لتعزيز الوعي المجتمعي حول القضايا المرتبطة بهذه التكنولوجيا وضمان استخدامها بطريقة مسئولة ومستدامة تعود بالنفع على الجميع.
- وأخيرًا، يجب إجراء المزيد من البحوث لاستكشاف المعايير الأخلاقية من منظور المطورين والمستخدمين، مما يساعد على فهم التحديات المرتبطة بها. كما ينبغي تقديم حلول عملية لتطبيق كل مبدأ أخلاقي بفعالية في أنظمة الذكاء الاصطناعي وذلك لتقليل الفجوة بين المبادئ النظرية والتطبيق العملي وتطوير أنظمة ذكية مسئولة.

قائمة المراجع

أولا: المراجع العربية:

- أحمد العربي (٢٠٢١)، الذكاء الاصطناعي: ثورة التكنولوجيا القادمة، مؤسسة الفكر العربي.
- الجهاز المركزي للتعبئة العامة والإحصاء (٢٠٢٢)، الاستراتيجية الوطنية للذكاء الاصطناعي في مصر: رؤية مستقبلية، القاهرة: الجهاز المركزي للتعبئة العامة والإحصاء.
- الجهاز المركزي للمحاسبات (٢٠٢١)، تقرير عن دور الذكاء الاصطناعي في مكافحة الفساد، القاهرة: الجهاز المركزي للمحاسبات.
- المجلس الوطني للذكاء الاصطناعي (٢٠٢٣)، تقرير عن التقدم المحرز في تنفيذ الاستراتيجية الوطنية للذكاء الاصطناعي، القاهرة: المجلس الوطني للذكاء الاصطناعي.
- هيئة الرقابة الإدارية (٢٠٢٣)، الاستراتيجية الوطنية لمكافحة الفساد. القاهرة: هيئة الرقابة الإدارية.
- هيئة الرقابة الإدارية (٢٠٢٣)، الرقابة الإدارية تضبط تشكيل عصابي تخصص في اصطناع وتزوير المستندات الحكومية. (https://aca.gov.eg/News/5217.aspx)
- وزارة الاتصالات وتكنولوجيا المعلومات (٢٠٢١)، الاستراتيجية الوطنية للذكاء الاصطناعي في مصر، القاهرة: وزارة الاتصالات وتكنولوجيا المعلومات.
- وزارة الاتصالات وتكنولوجيا المعلومات (٢٠٢٣)، الاستراتيجية الوطنية للذكاء الاصطناعي. القاهرة: وزارة الاتصالات وتكنولوجيا المعلومات.

ثانيا: المراجع الأجنبية:

- Agrawal, A., Gans, J., & Goldfarb, A. (2018). Prediction Machines: The Simple Economics of Artificial Intelligence. Harvard Business Review Press .
- Al-Garadi, M. A., Mohamed, A., Al-Ali, A. K., Du، X., Ali, I & ،.Guizani, M. (2020). A survey of machine and deep learning methods for internet of things (IoT) security. *IEEE Communications Surveys & Tutorials*, 22(3), 1646-1685.
- Anderson, M. (2007). *Machine ethics: Creating an ethical intelligent agent*. AI & Society, 22(4), 477-493.
- Anderson, M., & Anderson, S. (2007). *Machine ethics: New directions*. Oxford University Press.
- Berendt, B (2019) AI for the common good?! Pitfalls, challenges, and ethics pen-testing. *Paladyn, Journal of Behavioral Robotics* 10(1): 44–65. DOI: 10.1515/pjbr-2019-0004.
- Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- Brey, P., & Dainow, B. (2023). Ethics by design for artificial intelligence. *AI and Ethics*, 1-13.

- Brundage, L., Avin, S., Clark, J., Toner, M., Eckersley, P., Garfinkel, S., & Anderson, R. (2020). Toward trustworthy AI development: Aligning theoretical and practical perspectives. arXiv preprint arXiv:2004.07218.
- Brynjolfsson, E & McAfee, A. (2014). The Second Machine Age: Work Progress and Prosperity in a Time of Brilliant Technologies. W. W. Norton & Company .
- Bryson, J. (2018). The artificial intelligence of the ethics of artificial intelligence: An introductory overview for policymakers. *Cambridge Handbook of Artificial Intelligence*, 316-334.
- Bryson, J. J., Taddeo, M & Floridi, L. (2019). Ethics of artificial intelligence. The Oxford handbook of digital ethics، 109-139.
- Cath, C. (2018). *Governing artificial intelligence: Ethical, legal, and societal issues*. Philosophy & Technology, 31(4), 611-627.
- Cath, C; Wachter, S; Mittelstadt, B. (2018) Artificial intelligence and the 'good society': The US, EU, and UK approach. *Science and Engineering Ethics* 24(2): 505–528. DOI: 10.1007/s11948-017-9901-7.
- Craglia M, Annoni, A & Benczur. P. (2018). *Artificial intelligence : a European perspective*. Publications Office of the European Union, Luxembourg
- Doshi-Velez, F & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608.
- European Commission. (2020). *White paper on artificial intelligence: A European approach to excellence and trust*. Publications Office of the European Union.
- Flanagan, W. (2024). Ethics vs Laws: 5 Key Differences Experts Want You to Know in 2024. Available at: https://www.infonetica.net/articles/ethics-vs-laws
- Floridi, L. (2019). *The ethics of artificial intelligence: Principles, challenges, and opportunities*. Minds and Machines, 29(4), 569-589.
- Floridi, L. (2020). *The ethics of artificial intelligence*. Oxford University Press.
- Floridi, L; Cowls, J; Beltrametti, M. (2018). AI4People - an ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and machines* 28(4): 689–707.
- Ford, M. (2015). Rise of the Robots: Technology and the Threat of a Jobless Future. Basic Books.
- Goodman, B & Flaxman, S. (2017). European Union regulations on algorithmic decision-making and a "right to explanation". *AI Magazine*, 38(3), 50-57.
- Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines*, 30، 99-120.
- Harari, Y. (2016). *Homo Deus: A Brief History of Tomorrow*. Harper.
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389-399.
- Köbis, N; Starke, C & Rahwan, I. (2021). Artificial intelligence as an anti-corruption tool (AI-ACT). Potentials and pitfalls for top-down and bottom-up approaches.

- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- Machanavajjhala, a; Korolova, A & Sarma, A. (2011) Personalized social recommendations - accurate or private? Proceedings of the VLDB Endowment (PVLDB) 4(7).
- Marwick, A. E & Lewis, R. (2017). Media Manipulation and Disinformation Online. Data & Society Research Institute.
- McKendrick, J. (2022). 7 Steps to More Ethical Artificial Intelligence. Available at: https://www.forbes.com/sites/joemckendrick/2022/06/10/7-steps-to-more-ethical-artificial-intelligence/
- Mittelstadt, B. D. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence, 1*(11), 501-507.
- Mittelstadt, B. D., Allo, P., Taddeo, M &.Wachter, S. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 2053951716679679.
- Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From what to how: An initial review of publicly available AI ethics tools, methods, and research to translate principles into practices. *Science and Engineering Ethics, 26*(4), 2141-2168.
- Naqvi, A. (2020). *Artificial intelligence for audit، forensic accounting، and valuation: A strategic perspective*. John Wiley & Sons. https://doi.org/10.1002/978111960 1906
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366 (6464), 447-453 .
- OECD (2021). Recommendation of the council on artificial intelligence.
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group.
- Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J. F., Breazeal, C & .Rahwan, T. (2019). Machine behavior. *Nature*, 568(7753), 477-486.
- Russell, S., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson.
- Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 25, 1 pp. 1-65). Academic Press.
- Shin, D. & Park, Y. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior* 98: 277–284.
- Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. International. *Journal of Human–Computer Interaction* 36(6): 495–504.

- Sreenu, D. (2023). How can artificial intelligence accelerate scientific research? https://scientificnirvana.com/how-can-artificial-intelligence-accelerate-scientific-research
- Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., & Walsh, T. (2016). Artificial intelligence and life in 2030. One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, Stanford University, Stanford, CA.
- Taddeo, M., & Floridi, L. (2018). How AI can be a force for good. *Science*, 361(6404), 751-752
- UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence.
- Wardle, C & Derakhshan, C .(2017). Information Disorder: Toward an interdisciplinary framework for research and policymaking. Council of Europe report, *DGI* (2017) 9 .
- Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. Public Affairs.
- Zuiderveen, F. J., Möller, J., S., Fathaigh, R., Irion, K., Dobber, T, & Hildebrandt, M. (2018). Discrimination، artificial intelligence، and algorithmic decision-making. *Utrecht Law Review*, 14(2), 69-201791.

ملحق (١) الصورة المبدئية لقائمة المعايير الأخلاقية لتصميم وتطوير تطبيقات الذكاء الاصطناعي

المجال الأخلاقي	المؤشرات	المراجع
الشفافية	– وعي المستخدمين بكيفية عمل التطبيق واتخاذ القرارات، وأن يكونوا قادرين على فهم الأسباب الكامنة وراء هذه القرارات. – الإفصاح بوضوح لجميع الأطراف المعنية عن هدف النظام، وقدراته، وقيوده، وفوائده، ومخاطره، والقرارات التي يتخذها. – تصميم أنظمة الذكاء الاصطناعي بحيث يتمكن الأفراد من تدقيق أنشطتها أو الاستفسار عنها أو الطعن فيها أو طلب تعديلها. ويشمل ذلك العمليات التنظيمية التي تتيح للمشغلين تلقي وتقييم طلبات الأطراف الخارجية. – اتخاذ التدابير التي تُمكّن من تتبع نظام الذكاء الاصطناعي ومراقبته لضمان الشفافية والمساءلة. – اتخاذ قرارات بشأن القضايا الأخلاقية، مثل كيفية إزالة التحيز من مجموعة بيانات. – تسجيل عمليات وأدوات التطوير للقرارات الأخلاقية بحيث يكون من الممكن فهم كيفية الالتزام بالمعايير الأخلاقية؛ مما يمكن من إجراء عمليات التدقيق، أو الاعتراض على قرارات النظام، أو لتصحيح أي مشكلات أخلاقية تنشأ بعد نشره.	(O'Neil, 2016; Floridi, 2019; Brey& Dainow, 2023)
التوضيح	– إمكانية توضيح أسباب اتخاذ القرارات المختلفة من قبل المطور والنظام الذكي. – قابلية قرارات الذكاء الاصطناعي للتفسير وإمكانية توضيح أسباب اتخاذ النظام قرارًا معينًا. – إتاحة آلية شرح اتخاذ القرارات والبيانات وقابلية التفسير لما يتخذ من قرارات.	(Zuboff, 2019; Brey& Dainow 2023)
المساءلة	– توفير آليات للمساءلة في حالة وقوع أضرار نتيجة لاستخدام التطبيق، وإمكانية الحصول على تعويض في حالة الضرر. – إتاحة إشراف بشري على دورات التشغيل واتخاذ قرار النظام من قبل	(Floridi, 2019; Shneiderman, 2020; Russell & Norvig, 2021)

المجال الأخلاقي	المؤشرات	المراجع
	أنظمة الذكاء الاصطناعي. – تضمين عملية نشر نظام الذكاء الاصطناعي تقييمًا للمخاطر، مع وجود إجراءات للتخفيف من تلك المخاطر بعد النشر، بحيث تكون جاهزة منذ لحظة تشغيل النظام. – قابلية أنظمة الذكاء الاصطناعي للتدقيق من قبل جهات مستقلة. ولا يقتصر التدقيق على قرارات النظام فحسب، بل يشمل أيضًا العمليات والأدوات المستخدمة أثناء عملية التطوير.	
الدقة	– قيام المطورين بتدريب النظام الذكي بشكل مكثف بحيث يعمل النظام بدقة عالية جدًا. – تقليل الأخطاء إلى أدنى حد ممكن، بحيث تكاد تكون منعدمة.	(Zuiderveen, et al., 2018; Naqvi, 2020)
العدالة	– تصميم التطبيقات بطريقة تضمن عدم التمييز أو التحيز ضد أي فئة من الناس. – تجنب أنظمة الذكاء الاصطناعي التحيز الخوارزمي، بما في ذلك التحيز في بيانات الإدخال، والنمذجة، وتصميم الخوارزميات. – إتاحة أنظمة الذكاء الاصطناعي عالميا قدر الإمكان وحيثما كان ذلك مناسبًا، وتوفر نفس الوظائف والمزايا لجميع المستخدمين بغض النظر عن قدراتهم أو معتقداتهم أو تفضيلاتهم أو اهتماماتهم. – تصميم أنظمة الذكاء الاصطناعي، قدر الإمكان، بحيث تتجنب التأثيرات الاجتماعية السلبية على الفئات الاجتماعية، وخاصة الفئات المحمية.	(Anderson, 2007; O'Neil, 2016; Anderson & Berendt, 2019)
الخصوصية	– تصميم التطبيقات بطريقة تحمي خصوصية المستخدمين. – تأمين التطبيقات لأمن البيانات الشخصية وحماية البيانات من الوصول غير المصرح به.	(Anderson, 2007; Bostrom, 2014 & Anderson)
الاستقلالية الآلية	– تدريب المطورين للنظام الذكي بشكل مناسب. – قدرة النظام على العمل دون إشراف بشري.	(Cath, 2018; Brey& Dainow, 2023)
الإشراف	– ضمان وجود إشراف بشري على القرارات الحساسة أو المهمة.	(Obermeyer et al., 2019;

المجال الأخلاقي	المؤشرات	المراجع
البشري/ الوكالة البشرية	– تصميم أنظمة ذكاء اصطناعي تساعد البشر في اتخاذ القرارات بدلاً من استبدالهم تمامًا. – عدم تصميم أو استخدام نظام الذكاء الاصطناعي بطريقة تحرم الأفراد من القدرة على اتخاذ القرارات التي ينبغي أن يكونوا قادرين على اتخاذها بأنفسهم. – عدم تصميم أو استخدام الذكاء الاصطناعي بطريقة تؤدي إلى تقليص الحريات الأساسية للإنسان، بما في ذلك حرية التنقل والتجمع والتعبير والحصول على المعلومات. – تصميم أو استخدام أنظمة الذكاء الاصطناعي لخلق الإدمان على النظام أو على الخدمات التي يقدمها. – عدم اتخاذ أنظمة الذكاء الاصطناعي القرار النهائي بشأن القضايا المهمة ذات الطابع الشخصي أو الأخلاقي أو السياسي ولكن يمكنها تقديم التوصيات.	Brundage et al., 2020 & Brey& Dainow, 2023)
الأمن	– تصميم التطبيقات بطريقة تضمن سلامة المستخدمين والمجتمع، وعدم استخدامها في أغراض ضارة أو غير أخلاقية. – احترام حقوق أصحاب البيانات أثناء معالجة البيانات الشخصية. – توفير آليات توضح كيفية تحقيق الشرعية والعدالة والشفافية عند معالجة البيانات الشخصية بواسطة الذكاء الاصطناعي. – اتخاذ تدابير لحماية حقوق أصحاب البيانات من خلال إجراءات تقنية، مثل إخفاء الهوية، وكذلك إجراءات تنظيمية، مثل أنظمة التحكم في الوصول. – تدعيم أنظمة الذكاء الاصطناعي لحق الأفراد في سحب موافقتهم على استخدام بياناتهم الشخصية. – جمع البيانات وتخزينها ومعالجتها بطريقة تتيح تدقيقها من قبل البشر.	(Bostrom, 2014; Al-Garadi et al., 2020; Russell & Norvig, 2021)
الموثوقية	– حماية الأنظمة من الهجمات الإلكترونية والتلاعب.	(Craglia et al., 2018;

المجال الأخلاقي	المؤشرات	المراجع
والأمن السيبراني	– ضمان استمرارية الخدمة وعدم تعطلها بسبب التهديدات. – الحفاظ على أمان الخدمة لحماية المستخدمين والبيانات.	Brey& Dainow, 2023)
المسئولية الاجتماعية	– وعي المطورين والمنتجين بالآثار الاجتماعية لتطبيقاتهم. – الأخذ بعين الاعتبار التأثير المحتمل للتطبيقات على المجتمع. – السعي الدائم لتصميم تطبيقات تخدم الصالح العام.	(Anderson, & Anderson 2007; Russell & Norvig, 2021)
القيم البشرية	– احترام أنظمة الذكاء الاصطناعي حقوق الإنسان. – مراعاة التنوع والاختلاف بين الأفراد والمجتمعات. – ضمان استقلالية الأفراد في اتخاذ قراراتهم دون قيود غير مبررة.	(Köbis et al., 2021; Brey& Dainow, 2023)
الرفاهية البشرية والاجتماعية والبيئية	– إفادة أنظمة الذكاء الاصطناعي الأفراد والمجتمع والبيئة. – التأكد من أن استخدام الذكاء الاصطناعي يعزز حقوق الإنسان. – إسهام الذكاء الاصطناعي في تحقيق الرفاهية العامة. – منع استخدام الذكاء الاصطناعي بطريقة تقوض من حقوق الإنسان ورفاهيته.	(Brynjolfsson & McAfee,2014; Brey& Dainow, 2023)
الاستدامة والمسئولية البيئية	– ضمان أن تكون أنظمة الذكاء الاصطناعي آمنة للاستخدام وألا يكون لديها ميل لإلحاق الضرر أو التسبب في تدهور كبير في الصحة أو الرفاهية الجسدية أو النفسية لأي من الأطراف المعنية، بما في ذلك المستخدمين والعملاء وأصحاب البيانات والأطراف الأخرى المتأثرة. – مراعاة تطوير الذكاء الاصطناعي لمبادئ الاستدامة البيئية. – عدم تأثير أنظمة الذكاء الاصطناعي سلبًا على جودة التواصل أو التفاعل الاجتماعي أو تدفق المعلومات أو العلاقات الاجتماعية أو العمليات الديمقراطية.	(Jobin et al., 2019; Floridi, 2019; Brey& Dainow, 2023; Flanagan, 2024)

ملحق (٢) قائمة بأسماء السادة المحكمين

م	الاسم	الجامعة	الكلية	الوظيفة
١	إنجي عبد المعبود بيومي	مدينة السادات	الحاسبات والذكاء الاصطناعي	مدرس علوم الحاسب
٢	أحمد عبدالحميد صالح	دمنهور	الحاسبات والمعلومات	مدرس علوم الحاسب
٣	رشا محمد منتصر	دمنهور	الحاسبات والمعلومات	مدرس نظم المعلومات
٤	عمرو فتحي السيد	دمنهور	الحاسبات والمعلومات	مدرس علوم الحاسب
٥	مي حلمي شاهين	دمنهور	الحاسبات والمعلومات	مدرس تكنولوجيا المعلومات
٦	أحمد حامد عطية	دمنهور	الحاسبات والمعلومات	مدرس علوم الحاسب
٧	محمد عبداللطيف سيد الأهل	دمنهور	الحاسبات والمعلومات	مدرس علوم الحاسب

ملحق (٣) قائمة بتعديلات السادة المحكمين

المحكم الأول:

- إضافة مؤشر جديد تحت معيار الشفافية.
- إضافة مؤشر جديد تحت معيار الاستقلالية الآلية.
- إضافة مؤشر جديد تحت معيار الإشراف البشري.
- إضافة مؤشر جديد تحت معيار الرفاهية البشرية.
- إضافة معيار جديد (التكيف) وإضافة مؤشرين لهذا المعيار.

المحكم الثاني:

- تعديل صياغة المؤشرات (الأول، والثاني، والثالث، والخامس) تحت معيار الشفافية
- إضافة مؤشر جديد تحت معيار المحاسبية.
- إضافة جزء مكمل للمؤشر الأول في معيار الدقة.
- إضافة مؤشر جديد تحت معيار الدقة.
- تعديل صياغة المؤشر الخامس تحت مجال الإشراف البشري.
- إضافة مؤشر لمعيار المسئولية الاجتماعية.

المحكم الثالث:

- بمعيار الدقة مكتوب "تقليل الأخطاء إلى أدنى حد ممكن، بحيث تكاد تكون منعدمة". نقترح أن تحدد نسبة الخطأ المقبولة.
- هناك نقطتان ممكن أن يكون بينهما تضارب في معيار الاستقلال "قدرة النظام على العمل دون إشراف بشري" وفي الإشراف البشري "ضمان وجود الإشراف البشري على القرارات الحاسمة أو المهنية" ونقترح تعديلها في معيار الاستقلالية بحيث تكون: "قدرة النظام على العمل دون إشراف بشري ماعدا في القرارات الحاسمة أو المهنية".
- بمعيار الوكالة البشرية مكتوب "تصميم أو استخدام أنظمة الذكاء الاصطناعي لخلق الإدمان على النظام أو على الخدمات التي يقدمها". نقترح إعادة صياغتها على النحو التالي: عدم تصميم أو استخدام أنظمة الذكاء الاصطناعي لخلق الإدمان على النظام أو على الخدمات التي يقدمها.

المحكم الرابع:

- تعديل صياغة المؤشر الأول تحت معيار الشفافية.
- إعادة صياغة المؤشر الثالث تحت معيار الشفافية.
- إضافة مؤشر جديد تحت معيار التوضيح.
- إضافة مؤشر جديد تحت معيار الدقة.

- إضافة مؤشر جديد تحت معيار الخصوصية.
- إضافة مؤشر جديد تحت معيار الاستقلالية.
- إضافة مؤشر جديد تحت معيار الأمن.
- إضافة مؤشر جديد تحت معيار الموثوقية.
- إضافة مؤشر جديد تحت معيار المسئولية الاجتماعية.
- إضافة مؤشر جديد تحت معيار القيم البشرية.
- إضافة مؤشر جديد تحت معيار الرفاهية البشرية.
- إضافة مؤشرين تحت معيار الاستدامة والمسئولية البيئية.

المحكم الخامس:

- إضافة مؤشر جديد لمعيار الخصوصية.
- إعادة صياغة أحد مؤشرات معيار الاستقلالية.

المحكم السادس:

- إعادة صياغة مؤشرات معيار المساءلة.
- إعادة صياغة مؤشرات معيار العدالة.

المحكم السابع:

- إعادة صياغة مؤشرات معيار الأمن.
- إعادة صياغة مؤشرات معيار الاستدامة والمسئولية البيئية.

‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘‘

ملحق (٤) القائمة النهائية للمعايير الأخلاقية لتصميم وتطوير تطبيقات الذكاء الاصطناعي

المجال الأخلاقي	المؤشرات	المراجع
الشفافية	– رفع وعي المستخدمين بكيفية عمل التطبيق الذكي واتخاذ القرارات، وأن يكونوا قادرين على فهم الخدمات الذكية وأيضاً الأسباب الكامنة وراء هذه القرارات. – الإفصاح بوضوح لجميع الأطراف المعنية عن هدف النظام الذكي، وقدراته، وقيوده، وفوائده، ومخاطره، والقرارات التي يتخذها. – إمكانية تدقيق النظام والاستفسار عن قراراته أو الطعن فيها ويشمل ذلك العمليات التنظيمية التي تتيح للمستخدمين تلقي وتقييم الخدمات التي يقدمها النظام الذكي. – اتخاذ التدابير التي تُمكّن من تتبع نظام الذكاء الاصطناعي ومراقبته لضمان الشفافية والمساءلة. – اتخاذ قرارات بشأن القضايا الأخلاقية، مثل كيفية إزالة التحيز إلى فئة معينة من البيانات. – تسجيل عمليات وأدوات التطوير للقرارات الأخلاقية بحيث يكون من الممكن فهم كيفية الالتزام بالمعايير الأخلاقية؛ مما يمكن من إجراء عمليات التدقيق، أو الاعتراض على قرارات النظام، أو لتصحيح أي مشكلات أخلاقية تنشأ بعد نشره. – إعلام المستخدمين بشكل واضح عن أنواع البيانات التي يتم جمعها والغرض من جمعها.	(O'Neil, 2016; Floridi, 2019; Brey& Dainow, 2023)
التوضيح	– إمكانية توضيح أسباب اتخاذ القرارات المختلفة من قبل المطور والنظام الذكي – قابلية قرارات الذكاء الاصطناعي للتفسير وإمكانية توضيح	(Zuboff, 2019; Brey& Dainow 2023)

المجال الأخلاقي	المؤشرات	المراجع
	أسباب اتخاذ النظام لقرار معين. – إتاحة آلية شرح اتخاذ القرارات والبيانات وقابلية التفسير لما يتخذ من قرارات. – تطوير أنظمة تفسير آلية (Explainable AI - XAI) لضمان أن تكون قرارات الذكاء الاصطناعي قابلة للفهم من قبل الجميع.	
المساءلة	– توفير آليات للمساءلة في حالة وقوع أضرار نتيجة لاستخدام التطبيق، وإمكانية الحصول على تعويض في حالة الضرر. – إتاحة إشراف بشري على دورات التشغيل واتخاذ القرار النظام من قبل أنظمة الذكاء الاصطناعي. – تضمين عملية نشر نظام الذكاء الاصطناعي تقييمًا للمخاطر، مع وجود إجراءات للتخفيف من تلك المخاطر بعد النشر، بحيث تكون جاهزة منذ لحظة تشغيل النظام. – قابلية أنظمة الذكاء الاصطناعي للتدقيق من قبل جهات مستقلة. ولا يقتصر التدقيق على قرارات النظام فحسب، بل يشمل أيضًا العمليات والأدوات المستخدمة أثناء عملية التطوير. – المتابعة المستمرة للأنظمة الذكية لمتابعة حدوث مخاطر مستقبلية غير متوقعة.	(Floridi, 2019; Shneiderman, 2020; Russell & Norvig, 2021)
الدقة	– قيام المطورين بتدريب النظام الذكي بشكل مكثف بحيث يعمل النظام بدقة عالية جدًا. مما يتطلب تدريب النظام على بيانات ضخمة متعددة الأنماط. – استخدام تقنيات التعلم الآلي القابلة للتدقيق لضمان عدم وجود	(Zuiderveen, et al., 2018; Naqvi, 2020)

المجال الأخلاقي	المؤشرات	المراجع
	أخطاء غير مكتشفة. – تقليل الأخطاء إلى أدنى حد ممكن، بحيث لا تزيد عن نسبة ١٪ – استخدام مقاييس نماذج الذكاء الاصطناعي لقياس دقة النماذج التي تستخدمها التطبيقات.	
العدالة	– تصميم التطبيقات بطريقة تضمن عدم التمييز أو التحيز ضد أي فئة من الناس. – تجنب أنظمة الذكاء الاصطناعي التحيز الخوارزمي، بما في ذلك التحيز في بيانات الإدخال، والنمذجة، وتصميم الخوارزميات. – إتاحة أنظمة الذكاء الاصطناعي عالميا قدر الإمكان وحيثما كان ذلك مناسبًا، وتوفر نفس الوظائف والمزايا لجميع المستخدمين بغض النظر عن قدراتهم أو معتقداتهم أو تفضيلاتهم أو اهتماماتهم. – تصميم أنظمة الذكاء الاصطناعي، قدر الإمكان، بحيث تتجنب التأثيرات الاجتماعية السلبية على الفئات الاجتماعية، وخاصة الفئات المحمية.	(Anderson, 2007; O'Neil, 2016; Anderson & Berendt, 2019)
الخصوصية	– تصميم التطبيقات بطريقة تحمي خصوصية المستخدمين. – تأمين التطبيقات لأمن البيانات الشخصية وحماية البيانات من الوصول غير المصرح به. – استخدام تقنيات مثل التشفير والتجميع (Federated Learning) لحماية البيانات – الحق في النسيان (Right to be Forgotten) هو مبدأ	(Anderson, 2007; Bostrom, 2014 & Anderson)

المجال الأخلاقي	المؤشرات	المراجع
	قانوني يمنح الأفراد الحق في طلب حذف بياناتهم الشخصية من قواعد بيانات الشركات أو مزودي الخدمات الرقمية عندما لا يكون هناك سبب قانوني للاحتفاظ بها. هذا المفهوم جاء كجزء أساسي من اللائحة العامة لحماية البيانات (GDPR) في الاتحاد الأوروبي، وهو يعزز خصوصية الأفراد عبر الإنترنت.	
الاستقلالية الآلية	– تدريب المطورين للنظام الذكي بشكل مناسب. – قدرة النظام على العمل دون إشراف بشري ماعدا في القرارات الحاسمة أو المهنية. – ضمان أن يكون للمستخدمين السيطرة الكاملة على بياناتهم الشخصية. – إلزام المطورين بتوفير أدوات مراقبة تسمح بالتدخل البشري في حالات الطوارئ.	(Cath, 2018; Brey& Dainow, 2023)
الإشراف البشري/ الوكالة البشرية	– ضمان وجود إشراف بشري على القرارات الحساسة أو المهمة. – تصميم أنظمة ذكاء اصطناعي تساعد البشر في اتخاذ القرارات بدلاً من استبدالهم تمامًا. – تحديد قائمة بالقرارات الحساسة التي يجب أن تبقى تحت إشراف بشري كامل. – عدم تصميم أو استخدام نظام الذكاء الاصطناعي بطريقة تحرم الأفراد من القدرة على اتخاذ القرارات التي ينبغي أن يكونوا قادرين على اتخاذها بأنفسهم. – عدم تصميم أو استخدام الذكاء الاصطناعي بطريقة تؤدي إلى تقليص الحريات الأساسية للإنسان، بما في ذلك حرية التنقل والتجمع والتعبير والحصول على المعلومات.	(Obermeyer et al., 2019; Brundage et al., 2020 & Brey& Dainow, 2023)

المجال الأخلاقي	المؤشرات	المراجع
	– عدم تصميم أو استخدام أنظمة الذكاء الاصطناعي لخلق الإدمان على النظام أو على الخدمات التي يقدمها. – عدم اتخاذ أنظمة الذكاء الاصطناعي القرار النهائي بشأن القضايا المهمة ذات الطابع الشخصي أو الأخلاقي أو السياسي ولكن يمكنها تقديم التوصيات. – إجراء تقييمات دورية لضمان استمرارية الامتثال للمعايير الأخلاقية.	
الأمن	– تصميم التطبيقات بطريقة تضمن سلامة المستخدمين والمجتمع، وعدم استخدامها في أغراض ضارة أو غير أخلاقية. – احترام حقوق أصحاب البيانات أثناء معالجة البيانات الشخصية. – توفير آليات توضح كيفية تحقيق الشرعية والعدالة والشفافية عند معالجة البيانات الشخصية بواسطة الذكاء الاصطناعي. – اتخاذ تدابير لحماية حقوق أصحاب البيانات من خلال إجراءات تقنية، مثل إخفاء الهوية، وكذلك إجراءات تنظيمية، مثل أنظمة التحكم في الوصول. – تدعيم أنظمة الذكاء الاصطناعي لحق الأفراد في سحب موافقتهم على استخدام بياناتهم الشخصية. – جمع البيانات وتخزينها ومعالجتها بطريقة تتيح تدقيقها من قبل البشر. – توفير آليات استجابة سريعة في حالة اكتشاف ثغرات أمنية.	(Bostrom, 2014; Al-Garadi et al., 2020; Russell & Norvig, 2021)
الموثوقية والأمن السيبراني	– حماية الأنظمة من الهجمات الإلكترونية والتلاعب. – ضمان استمرارية الخدمة وعدم تعطلها بسبب التهديدات.	(Craglia et al., 2018; Brey& Dainow, 2023)

المجال الأخلاقي	المؤشرات	المراجع
	– الحفاظ على أمان الخدمة لحماية المستخدمين والبيانات. – إلزام الشركات بتوفير خطط طوارئ في حالة تعطل الخدمة.	
المسئولية الاجتماعية	– وعي المطورين والمنتجين بالآثار الاجتماعية لتطبيقاتهم. – الأخذ بعين الاعتبار التأثير المحتمل للتطبيقات على المجتمع. – السعي الدائم لتصميم تطبيقات تخدم الصالح العام. – التوعية بمميزات وأضرار الذكاء الاصطناعي. – توفير آليات لتلقي ملاحظات المستخدمين وتحسين الأنظمة بناءً عليها.	(Anderson, & Anderson 2007; Russell & Norvig, 2021)
القيم البشرية	– احترام أنظمة الذكاء الاصطناعي حقوق الإنسان. – مراعاة التنوع والاختلاف بين الأفراد والمجتمعات. – ضمان استقلالية الأفراد في اتخاذ قراراتهم دون قيود غير مبررة. – توفير تدريب للمطورين حول كيفية تصميم أنظمة تحترم القيم الإنسانية.	(Köbis et al., 2021; Brey& Dainow, 2023)
الرفاهية البشرية والاجتماعية والبيئية	– إفادة أنظمة الذكاء الاصطناعي الأفراد والمجتمع والبيئة. – التأكد من أن استخدام الذكاء الاصطناعي يعزز حقوق الإنسان. – إسهام الذكاء الاصطناعي في تحقيق الرفاهية العامة. – منع استخدام الذكاء الاصطناعي بطريقة تقوض من حقوق الإنسان ورفاهيته. – تصميم أنظمة الذكاء الاصطناعي بحيث تكون قابلة للاستخدام من قبل الأشخاص ذوي الإعاقات أو الفئات	(Brynjolfsson & McAfee,2014; Brey& Dainow, 2023)

المجال الأخلاقي	المؤشرات	المراجع
	المهمشة. – إلزام الشركات بإجراء تقييمات للأثر البيئي لأنظمتها.	
الاستدامة والمسئولية البيئية	– ضمان أن تكون أنظمة الذكاء الاصطناعي آمنة للاستخدام وألا يكون لديها ميل لإلحاق الضرر أو التسبب في تدهور كبير في الصحة أو الرفاهية الجسدية أو النفسية لأي من الأطراف المعنية، بما في ذلك المستخدمين والعملاء وأصحاب البيانات والأطراف الأخرى المتأثرة. – مراعاة تطوير الذكاء الاصطناعي لمبادئ الاستدامة البيئية. – عدم تأثير أنظمة الذكاء الاصطناعي سلبًا على جودة التواصل أو التفاعل الاجتماعي أو تدفق المعلومات أو العلاقات الاجتماعية أو العمليات الديمقراطية. – تطوير معايير لضمان أن تكون أنظمة الذكاء الاصطناعي صديقة للبيئة. – تشجيع استخدام الطاقة المتجددة في تشغيل مراكز البيانات.	(Jobin et al., 2019; Floridi, 2019; Brey& Dainow, 2023; Flanagan, 2024)
التكيف	– القدرة على التكيف مع التغيرات – ضمان قدرة النظام على التكيف مع السياقات المختلفة دون فقدان الفعالية أو الدقة.	Salehie, M., & Tahvildari, L. (2009). (TAAS)