

الاسم: عبدالرحمن الميمان الرقم الجامعي: 441170135



في هذه الأيام و مع التطور السريع لتقنيات الذكاء الاصطناعي واستخداماتها في مختلف القطاعات مثل الصحة والتعليم، وازدياد الاعتماد عليها لتحسين كفاءة العمليات صنع القرار، أصبح من الضروري وضع إطار أخلاقي يضمن استخدام هذه التقنيات بشكل يراعي القيم الإنسانية ويحقق التنمية المستدامة. لذلك، طوّرت الهيئة السعودية للبيانات والذكاء الاصطناعي (SDAIA) مبادئ توجيهية تهدف لدعم الابتكار وضمان سلامة المجتمع، مع الالتزام بالمعايير الأخلاقية.

تأتي هذه المبادئ مواكبةً للتشريعات والقوانين العالمية، لتتماشى مع توجهات الدول المتقدمة في تنظيم تقنيات الذكاء الاصطناعي، وتضمن توافق الأنظمة المستخدمة في المملكة مع أفضل الممارسات الدولية، بما يعزز الثقة في استخدامها على النطاق الوطني والدولي.



المبدأ الأول - النزاهة والإنصاف:
يتطلب هذا المبدأ تصميم أنظمة الذكاء الاصطناعي بطريقة تضمن الشمولية وتجنب التحيز والتمييز. يحرص المطورون على تنقية وتنظيف البيانات المستخدمة في تدريب الأنظمة من أي تحيز، ويعتمدون على مصادر متنوعة لتمثيل شرائح المجتمع كافة. يشمل ذلك تجنب الاعتماد على عوامل مثل العرق أو الجنس أو العمر أو الدين في اتخاذ القرارات، وضمان أن تكون مخرجات النظام عادلة للجميع.



المبدأ الثاني - الخصوصية والأمن:
يشدد هذا المبدأ على حماية خصوصية الأفراد وأمن بياناتهم. يتطلب من المطورين تصميم أنظمة تحافظ على سرية البيانات الشخصية وتعتمد بروتوكولات صارمة للحد من الوصول غير المصرح به. كما يشمل توفير آليات لمعالجة البيانات بطريقة آمنة، مثل التشفير وإخفاء الهوية، بالإضافة إلى الالتزام بالمعايير السيبرانية الدولية لضمان سلامة النظام من الهجمات.



المبدأ الثالث - الإنسانية:
يعزز هذا المبدأ القيم الإنسانية والثقافية الأساسية، ويؤكد على استخدام الذكاء الاصطناعي لتحسين حياة الأفراد بدلاً من تقييدها. يجب أن تساهم هذه الأنظمة في دعم قدرات الأفراد وتمكينهم بدلاً من التلاعب بهم. على سبيل المثال، يجب أن تكون أدوات الذكاء الاصطناعي وسيلة لتحقيق الأهداف الإيجابية للإنسانية وتعزيز الابتكار والمسؤولية الاجتماعية.



المبدأ الرابع - المنافع الاجتماعية والبيئية:
يسعى هذا المبدأ إلى جعل تقنيات الذكاء الاصطناعي ذات تأثير إيجابي على المجتمع والبيئة. يتطلب ذلك أن يتم توجيه هذه الأنظمة لدعم الأهداف الاجتماعية، مثل تحسين جودة الحياة، ودعم الاستدامة البيئية، ومكافحة التحديات البيئية كالتغير المناخي. على المطورين التأكد من أن الأنظمة لا تسبب أضراراً أو تهدد البيئة، بل تساهم في تعزيز التنمية المستدامة.



المبدأ الخامس - الموثوقية والسلامة:
يركز هذا المبدأ على تطوير أنظمة ذكاء اصطناعي تتمتع بالموثوقية وقادرة على الأداء بفعالية تحت مختلف الظروف. يجب على المطورين إنشاء أنظمة قوية يمكن الاعتماد عليها حتى في الحالات الطارئة، وذلك من خلال ضمان سلامة الخوارزميات وتوفير آليات تمنع حدوث ضرر غير متوقع ولا يشكل خطر على حياة الأفراد. كما يتطلب هذا المبدأ أن تكون الأنظمة سهلة الاختبار والتقييم الدوري لضمان استمرار جودتها.



المبدأ السادس - الشفافية والقابلية للتفسير:
يشدد هذا المبدأ على ضرورة أن تكون الأنظمة شفافة وسهلة التفسير، بحيث تتيح للمستخدمين فهم كيفية عملها وطريقة اتخاذها للقرارات. يتطلب ذلك من المطورين بناء آليات تتبع للقرارات المؤتمتة وتوثيق الخطوات المتعلقة بتدريب النظام. هذا يعزز من ثقة المستخدمين وبتيح لهم إمكانية التحقق من مخرجات النظام وتحليلها، مما يسهم في فهم أوسع لأثر النظام واستخداماته ويكون مطورها قادر على التبرير أسس تصميمها وممارساتها وعملياتها وخوارزمياتها وقراراتها وسلوكياتها المسموح بها أخلاقيا وغير الضارة للعامّة.



المبدأ السابع - المساءلة والمسؤولية:
يهدف هذا المبدأ إلى تحميل مطوري ومشرفي الأنظمة الذكية المسؤولية عن تأثيراتها على المجتمع والأفراد. يجب عليهم تقديم إشراف بشري وحوكمة للتقليل من الأضرار المحتملة، ووضع آليات رقابية لضمان الاستخدام الأخلاقي للنظام. كما يشمل هذا المبدأ إمكانية مساءلة المسؤولين في حال حدوث أي تأثير سلبي غير مقصود أو إساءة في استخدام النظام، مما يعزز الالتزام بأعلى المعايير الأخلاقية.