CITRA

الهيئة العامة للاتصالات و تقنية المعلومات

COMMUNICATION & INFORMATION TECHNOLOGY REGULATORY AUTHORITY

الدليل الارشادي لأخلاقيات الذكاء الاصطناعي

V 1.1

المحتويات:

1. التعريفات 1
2. مقدمة: 2
3. نطاق التطبيق 3
4. الأهداف 3
5. قيم ومبادئ أخلاقيات الذكاء الاصطناعي 3
6. الخاتمة 5

1. التعريفات:

يكون للكلمات والعبارات التالية حيثما وردت في الدليل المعاني المخصصة لها أدناه، وتعتمد التعريفات الواردة في قانون الهيئة العامة للاتصالات وتقنية المعلومات رقم 2014/37 والمعدل بالقانون رقم 98 لسنة 2015 ولائحته التنفيذية ولائحة مصطلحات وتعريفات تقنية المعلومات والاتصالات الصادرة عن الهيئة:

1.1. الدولة: دولة الكويت.

1.2. الهيئة: الهيئة العامة للاتصالات وتقنية المعلومات.

1.3. أنظمة الذكاء الاصطناعي: الانظمة القادرة على أن تطبق التحليل المتقدم و التقنيات القائمة على المنطق، بما في ذلك التعلم الالي، لتفسير الاحداث و دعم القرارات و تطبيقها و اتخاذ القرارات ، وعلى سبيل المثال لا الحصر على ذلك الروبوتات وأنظمة المحادثة الآلية (chat bots) ، وأنظمة المساعدة الافتراضية الذكية) مثل (Alexa, Siri, Cortana, Google Assistant والسيارات ذاتية القيادة والطائرات المسيرة (drones) والمنازل الذكية والمدن الذكية وأنظمة توليد وعرض واكتشاف المحتوى وبعض تطبيقات الأجهزة الذكية. وتتنوع مجالات تطبيقات الذكاء الاصطناعي لتشمل التعليم والصحة والثقافة والتجارة والصناعة والنقل والبيئة والزراعة والأمن والترفيه وغيرها. ويشمل الذكاء الاصطناعي كلا من التعلم الآلي والتعلم العميق والتعلم المعزز، وقد حصلت عدة تطورات في الفترة الأخيرة في مجال الرؤية الحاسوبية ومعالجة وفهم اللغات الطبيعية باستخدام التعلم العميق، إضافة إلى إمكانية تفسير وفهم الخوارزميات ذات الصلة.

1.4. الجهة / الجهات: تشمل الوزارات والإدارات والمؤسسات والهيئات والشركات التابعة والمستقلة لحكومة دولة الكويت وشركات القطاع الخاص.

2. مقدمة:

في ظل التطورات المتسارعة للثورة الصناعية الرابعة وتوفر البيانات الرقمية وانخفاض تكاليف البنية التحتية والقدرات السحابية والتخزينية ونقل وتبادل البيانات والتعامل مع البيانات الضخمة وتمكين التحول الرقمي في كافة المجالات، أصبح استخدام تقنيات الذكاء الاصطناعي من الركائز لأساسية للتنمية الاقتصادية والاجتماعية ووسيلة فاعلة لتحقيق أهداف التنمية المستدامة وتمكين التحول الرقمي في الدولة.

كما أظهر استخدام الذكاء الاصطناعي دورا فاعلا لتطوير أنظمة وبرامج قادرة على محاكاة القدرات الذهنية للبشر منها التعلم والاستدلال والإبداع والتواصل وغيرها، وهذا التسارع في استخدام التطبيقات القائمة على تقنيات الذكاء الاصطناعي أدى إلى وجود تحديات اقتصادية واجتماعية تتعلق بالمسؤولية الأخلاقية والمخاطر الاجتماعية والتحولات الاقتصادية ومستقبل العمل وتعميق الفجوات في الفرص المتاحة لفئات معينة للعمل في هذا المجال أو حتى الاستفادة منه، وهذا الانتشار باستخدام الذكاء الاصطناعي والذي من المتوقع أن يتزايد بالسنوات القادمة قد يؤدي الى مخاطر وتهديدات تؤثر على الحكومات والاقتصادات والمجتمعات محليا و اقليميا و دوليا. أظهر استخدام الذكاء الاصطناعي دورا فاعلا لتطوير أنظمة وبرامج قادرة على محاكاة القدرات الذهنية للبشر منها التعلم والاستدلال والإبداع والتواصل وغيرها.

وبناء على ما تقدم ،قامت الهيئة العامة للاتصالات تقنية المعلومات بإعداد هذا الدليل الارشادي ليسهم في احترام وتعزيز ونشر حماية القيم والمبادئ والمعايير الأخلاقية فيما يخص استخدام تقنيات الذكاء الاصطناعي وعليه يمكن استخدام هذا الدليل والانتفاع به وتطبيقه لتحقيق الذكاء الاصطناعي الجدير بالثقة والموائم لتوصيات الأمم المتحدة الخاصة بأخلاقيات الذكاء الاصطناعي.

3. نطاق التطبيق.

يمكن لجميع الافراد والجهات التي تستخدم تقنيات الذكاء الاصطناعي في الدولة الاستفادة مما جاء في هذا الدليل الارشادي.

4. الأهداف:

4.1. يدعو هذا الدليل الارشادي الجهات والأفراد في الدولة إلى الالتزام الأخلاقي والمجتمعي والقانوني بالبنود المبينة أدناه، ذلك بغرض تعزيز الثقة بأنظمة الذكاء الاصطناعي وزيادة الاعتماد عليها.

4.2. يهدف هذا الدليل إلى تنظيم كافة الأنشطة والتطبيقات والمنتجات والخدمات والإجراءات التي تستند على الذكاء الاصطناعي، جزئيا أو كليا، أيا كان مصدرها (داخل أو خارج الدولة) ما دامت تؤثر على كل شخص طبيعي أو اعتباري داخل الدولة وتعتبر هذه البنود بوصلة أخلاقية تحكم تصرفات كافة الجهات التي تنظم وتستخدم وتصنع وتتعامل بأي شكل من الأشكال مع أنظمة الذكاء الاصطناعي والتأكد من انسجام كافة الأنشطة ذات العلاقة معها.

5. قيم ومبادئ أخلاقيات الذكاء الاصطناعي:

تم إنشاء الدليل الارشادي لأخلاقيات الذكاء الاصطناعي بالارتكاز على القيم والمبادئ لتعزيز تكنولوجيا الذكاء الاصطناعي في دولة الكويت على كافة الأصعدة التقنية والاجتماعية والسياسية، وتحدد هذه القيم والمبادئ إطاراً عاماً لتطوير أنظمة الذكاء الاصطناعي. وعليه، يجب احترام القيم والمبادئ المذكورة في هذه الوثيقة وتعزيزها عند الضرورة من خلال إعداد تشريعات ونظم ومبادئ توجيهية جديدة واجراء التعديلات بما يتوافق مع التشريعات والقوانين المعمول بها في دولة الكويت.

5.1. القيم:

5.1.1. احترام وحماية الأفراد من انتهاكات الذكاء الاصطناعي

ينبغي تطوير أنظمة الذكاء الاصطناعي بطريقة تحترم وتخدم وتحمي السلامة الجسدية والعقلية للبشر واحساسهم بالهوية وتلبى احتياجاتهم الأساسية، وتمنح البشر الذين يتفاعلون مع أنظمة الذكاء الاصطناعي القدرة على الحفاظ على تقرير المصير الكامل والفعال لأنفسهم بما في ذلك اختيار وتحديد رغباتهم الخاصة مع ضمان عدم وجود أي ضرر يلحق بالأفراد ومراعاة التقيد عند استخدام الذكاء الاصطناعي.

5.1.2. احترام الشريعة الاسلامية والدستور

ينبغي الحفاظ على مبادئ الشريعة الإسلامية والدستور والتشريعات وقوانين الدولة، وكذلك المعايير والإجراءات التي ينبغي أن تدعم هذه الأنظمة، لضمان حفظ الاجراءات القانونية الواجبة والمساواة أمام القانون بين جميع أفراد المجتمع.

5.1.3. حماية البيئة و تعزيز الاستدامة

مراعات المخاطر البيئية المتولدة من تشغيل وتدريب أنظمة الذكاء الاصطناعي من خلال تنمية الاستراتيجيات التي تحد من الملوثات البيئية، كما يتم استخدام أحدث التطورات التكنولوجية بهذا الشأن والتي من شأنها المحافظة على البيئة والتأكد من عدم تدهور النظم البيئية وبالتالي ضمان استدامتها للأجيال القادمة.

5.1.4. تعزيز الاستخدام السلمي لتحقيق رفاهية الافراد

يجب أن تعمل أنظمة الذكاء الاصطناعي على تحسين سبل الحياة للأفراد ويجب التحقق من أن تطبيق هذه الأنظمة لا تستخدم في مجالات تهدد تعايش البشر أو امنهم أو سلامتهم. كما يوصي هذا الدليل بأن يحظى الأشخاص الضعفاء من ذوي الإعاقة أو كبار السن أو غيرهم باهتمام أكبر وأن يتم تضمين احتياجاتهم في تطوير أنظمة الذكاء الاصطناعي.

5.2. المبادئ:

5.2.1. الشفافية وعدالة الإجراءات وقابلية التفسير:

مراعاة أن تكون أنظمة الذكاء الاصطناعي ومطوروها قادرين على تبرير أسس تصميمها وممارساتها وعملياتها وخوارزمياتها وقراراتها وسلوكياتها المسموح بها أخلاقياً وغير الضارة للعامة كما يجب أن تكون هذه الأنظمة قابلة للتتبع والتدقيق لضمان وتعزيز مبدأ الشفافية والقابلية للتفسير وعدالة الإجراءات لبناء الثقة في أنظمة وتقنيات الذكاء الاصطناعي.

ولضمان العدالة بكافة صورها أثناء استخدام الذكاء الاصطناعي، ينبغي إشراك جميع الأطراف ذات الصلة طوال دورة حياة نظم الذكاء الاصطناعي وتمكين الجميع في تطويرها بالإضافة إلى تمكين الأفراد والمجتمعات من الحصول على منافع تطبيقات الذكاء الاصطناعي بشكل متساو وعادل.

5.2.2. حفظ الحقوق والمساءلة:

وضع الآليات اللازمة لضمان تحمل المسؤولية وحفظ الحقوق وذلك عن طريق التركيز على وعي الأفراد ومسؤولياتهم الأخلاقية عن الممارسات في البيئة الافتراضية وعليه ينبغي تعزيز الوعي والدراية ونشر المعرفة الخاصة بوسائل تكنولوجيا الذكاء الاصطناعي واخلاقياته. كما يتطلب التأكيد على مسؤولية ضمان حق التقاضي والمحاسبة والمراجعة والاعتراض على الصعيد الفردي والجماعي، على مخرجات الذكاء الاصطناعي وطلب التفسير والتعويض عند التعدي على الحقوق وضمان الاستعانة بآراء الخبراء في الذكاء الاصطناعي في جلسات التقاضي المتعلقة بذلك، وتطوير تشريعات تضمن أجراء التدقيق الفني على أداء النظام وتقييم الخوارزميات والبيانات وعمليات التصميم واعداد التقارير للإسهام في تحسين موثوقية وسائل تكنولوجيا المعلومات.

5.2.3. حماية البيانات وحماية الخصوصية:

تطبيق حوكمة استخدام البيانات التي تحفظ جودة وسلامة البيانات والقدرة على معالجتها بطريقة تحمي خصوصية أصحاب البيانات الشخصية، والتأكد من تطبيق معايير الأمن السيبراني ذات العلاقة، بهدف منع الوصول غير المشروع إلى البيانات والنظام مما قد يؤدي إلى الإضرار بالسمعة، أو الأضرار النفسية، أو المالية أو المهنية.

5.2.4. الشمول والتنوع والحيادية والموضوعية وعدم التحيز:

ضمان تكافؤ الفرص واستفادة كافة الفئات والمناطق بشكل متساو وعادل من إمكانيات وفرص الذكاء الاصطناعي. كما ينبغي عدم التمييز بين الأشخاص والمجموعات والكيانات الإنسانية على أي أساس ديموغرافي أو صحي أو اقتصادي أو وضع سياسي أو نشاط قانوني وفقاً للتشريعات والقوانين المعمول بها في الدولة السارية، وضمان شمولية التطبيقات والخدمات لكافة القطاعات والأنشطة كالصحة والتعليم والاقتصاد والإعلام والثقافة. حيث ينبغي التأكد من إشراك كافة فئات المجتمع في التخطيط

والنقاش حول الذكاء الاصطناعي في كافة المراحل وعدم التقليل من الآراء بناء على اعتبارات غير علمية.

6. الخاتمة:

صمم هذا الدليل الارشادي باستخدام أفضل المعايير العالمية التي وردت في مجال اخلاقيات الذكاء الاصطناعي من منطلق المقاربة الإنسانية خاصة التوصيات التي وردت من منظمات عالمية كاليونيسكو، أدبيات IEEE وOECD، الاتحاد الأوروبي، الاتحاد الدولي للاتصالات، جامعة الدول العربية، مجلس التعاون الخليجي، ومصادر دولية وإقليمية أخرى. ويرتكز هذا الدليل على أسس القيم والكرامة الإنسانية والحريات الأساسية واستقلالية البشر والشمول والتنوع والعدالة والإنصاف والشفافية والمسائلة وحفظ الحقوق.