

أهم ما تضمنته مبادئ وإرشادات أخلاقيات الذكاء الاصطناعي من دبي الذكية

يجب أن تكون أنظمة الذكاء الاصطناعي عادلة وتطبق الشفافية وخاضعة للمساءلة وقابلة للشرح.

بقلم إم آي تي تكنولوجي ريفيو 09 يناير 2019

في ظل التطورات السريعة التي يمرُّ بها الذكاء الاصطناعي وتطبيقاته، يطالب الكثير من الخبراء وصانعي السياسات بضرورة وضع حدود له من خلال قواعد تنظم عمله وتضمن أن لا يؤدي استخدامه في القطاعات المختلفة إلى عواقب سلبية على البشر أو الأنظمة الأخرى. ولكن يبقى السؤال عن من سيكون المسؤول عن وضعها، وكيف نضمن من أنها ستكون عادلة ومنصفة، وفي نفس الوقت لا تقف كعائق في وجه تطور الذكاء الاصطناعي والابتكارات التي تتطور بشكل يومي. الأمر ليس سهلاً بالتأكيد، وقد ناقشنا سابقاً في مقالة [“هل نستطيع وضع قواعد أخلاقية للذكاء الاصطناعي”](#) الجهود المختلفة في وضع هذه القوانين والقواعد وجدواها. إضافة إلى لعبة [“الآلة الأخلاقية”](#) من جامعة إم آي تي التي تعد واحدة من أضخم الدراسات العالمية حول الأخلاقيات التي شارك فيها أشخاص من 233 بلد وسجلوا أكثر من 40 مليون قرار حول من على السيارة أن تدهس: الطفل أم الجدة.

دبي الذكية تطلق مبادئ وإرشادات أخلاقيات الذكاء الاصطناعي

وقد انضمت دبي، بوصفها المدينة العربية الأولى إلى هذه [الجهود](#)، حيث أعلن مكتب “دبي الذكية” إطلاق “مبادئ وإرشادات أخلاقيات الذكاء الاصطناعي” والتي تهدف المبادئ إلى تحقيق الاستخدام العادل لأنظمة الذكاء الاصطناعي، وجعلها قابلة للمساءلة، وجعلها قابلة للشرح تقنياً قدر الإمكان، وتحقيق الشفافية فيها. وتبدأ الوثيقة بشرح مبادئ دبي الأربعة للذكاء الاصطناعي هي الأخلاقيات، الأمان، الشمولية، البشرية. حيث أن “الأخلاقيات” تعني كون أنظمة الذكاء الاصطناعي عادلة وتطبق الشفافية وخاضعة للمساءلة وقابلة للفهم. أما “الأمان” فيشمل ضمان كونها آمنة ومسخرة لخدمة وحماية الإنسانية. ويقصد بـ “الشمولية” أن ينفع الذكاء الاصطناعي كافة أفراد المجتمع مع احترام كرامتهم وحقوقهم. وأخيراً “البشرية” والتي تشمل كون الذكاء الاصطناعي نافعا للبشرية وانسجامه مع القيم الإنسانية، في المدى القصير وكذلك البعيد.

ستقوم بتقديم الإرشادات لاستخدام تطبيقات الذكاء الاصطناعي، والتي تعزز من تحقيق قيم العدالة والشفافية والمساءلة لدى الأفراد والمؤسسات المعنية بتقديم خدمات الذكاء الاصطناعي في دبي. والتي ستؤدي بدورها إلى الارتقاء بالخدمات الحكومية وتعزيز فرص تبنيتها لحلول الذكاء الاصطناعي.

وبحسب البيان الصحفي: “تعالج مبادئ وإرشادات أخلاقيات الذكاء الاصطناعي قضايا مختلفة منها التحول السريع لواقع الذكاء الاصطناعي الذي أدى إلى نهج مجزأ في الأخلاقيات، بحيث تتعامل كل شركة مع القضايا الأخلاقية بطريقتها الخاصة. وتسعى هذه المبادئ إلى توضيح الغموض المحيط بمفهوم الأخلاق في الذكاء الاصطناعي؛ لا سيما وأن هذا الغموض قد يؤدي - كما يعتقد البعض - إلى كبت الابتكار داخل المنشآت، التي تختار وقف جهودها البحثية تحسباً للإجراءات الحكومية المستقبلية”.

أمثلة عملية عن حالات محتملة

الجدير بالذكر أن وثيقة المبادئ فيها الكثير من الأمثلة التي تشرح سيناريوهات لحالات محتملة لفهم أحد البنود. مثلاً جاء المثال التالي عند شرح إتاحة الفرصة للأشخاص المشاركين والمتأثرين بأنشطة الذكاء الاصطناعي الاعتراض على القرارات المؤتمتة الهامة المتعلقة بهم، واختيار عدم المشاركة عند الإمكان. حيث يجب أن تراعي المؤسسات توظيف مقيمين بشريين لمراجعة أي تحديات أثناء عملية اتخاذ القرارات، وتتيح فرصة رفض القرار المُعترض عليه عند اللزوم.

يسمح بنك من البنوك لعملائه بالتقدم للحصول على قرض عبر الإنترنت عبر إدخال بياناتهم. ويستخدم البنك نظاماً للذكاء الاصطناعي ليحدد آلياً قرار منح القرض ونسبة الفائدة التي سيفرضها. ويعطي البنك مستخدمي الخدمة خيار الاعتراض على القرار وطلب مراجعته من قبل موظف بشري⁷. ويطلب البنك من العميل إيضاح أسباب اعتراضه على القرار عن طريق ملء نموذج مخصص لهذا الغرض، مما يساعد الشخص المسؤول عن تدقيق الحالة ويقلل فرصة اعتراض العملاء على أي قرار دون أسباب وجيهة.

مصدر الصورة: وثيقة مبادئ وإرشادات أخلاقيات الذكاء الاصطناعي، موقع دبي الذكية الإلكتروني.

مراحل عدة وشركاء عالميون:

وقد تم بناء المبادئ والإرشادات في دبي الذكية باستخدام مقارنات معيارية دولية، وبالتواصل وجمع آراء الشركاء الاستراتيجيين من الجهات الحكومية الاتحادية والمحلية مثل هيئة تنظيم الاتصالات، مركز دبي للأمن الإلكتروني، هيئة الصحة في دبي، هيئة الطرق والمواصلات، بلدية دبي، هيئة كهرباء ومياه دبي، دائرة الأراضي والأملاك، وشركة اتصالات، بالإضافة إلى كبرى شركات التكنولوجيا بما في ذلك مايكروسوفت، آي بي إم، جوجل، وبي دبليو سي. وقد مرت هذه المبادئ بعدة إصدارات ومسودات حتى وصلت إلى الإطلاق أمس، في 8 يناير 2019.

الإصدار	المرحلة	تاريخ الإتمام	ملخص التغييرات
1.0	داخلية	2018.9.6	المسودة الأولى
1.1	داخلية	2018.9.6	تحسين
1.2	استشارات	2018.9.10	إضافة شرح تحت كل مبدأ
1.3	استشارات	2018.9.10	التغييرات عقب الاجتماع التوجيهي الثاني: المواءمة مع البشرية
1.4	استشارات	2018.9.25	التغييرات عقب الاجتماع التوجيهي الثاني: المساواة مع الشمولية
1.5	آراء وملاحظات	2018.10.9	التغييرات عقب الجولة الأولى من الآراء والملاحظات
1.6	مراجعة داخلية	2018.12.30	تعديلات ومراجعة لصياغة بعض أجزاء المحتوى وإضافة المقدمة

مصدر الصورة: وثيقة مبادئ وإرشادات أخلاقيات الذكاء الاصطناعي، موقع دبي الذكية الإلكتروني.

تقييم ذاتي

ومن الإعلانات المثيرة للاهتمام أيضاً إطلاق دبي الذكية لأول أداة تقييم ذاتي معتمدة في مجال استخدام أخلاقيات الذكاء الاصطناعي، والتي، بحسب البيان الصحفي، ستمكّن المطورين والعاملين في هذا المجال من تقييم امتثال مشاريعهم وتجاربهم في مجال الذكاء الاصطناعي للمبادئ التي تم وضعها.

لقراءة التقرير كاملاً: [مبادئ وإرشادات أخلاقيات الذكاء الاصطناعي](#).