“سدايا” تسخر تقنيات الذكاء الاصطناعي لخدمة اللغة العربية محليًا وعالميًا

29 أكتوبر، 2024

سخرت الهيئة السعودية للبيانات والذكاء الاصطناعي “سدايا” تقنيات الذكاء الاصطناعي لدعم اللغة العربية من خلال مجموعة من المبادرات التي تهدف إلى تعزيز مكانتها عالمياً وإثراء المحتوى العربي بموارد معرفية متقدمة ضمن إطار مستهدفات رؤية المملكة الرامية إلى دعم مكانة الريادية في خدمة اللغة العربية محلياً وإقليمياً وعالمياً.

وعكست هذه الجهود تتويج سدايا اليوم بجائزة مجمع الملك سلمان العالمي للغة العربية في دورتها الثالثة، لعام 2024م عن فئة المؤسسات: وذلك لتميّز إسهاماتها في خدمة اللغة العربية من خلال تطوير نماذج لغوية ضخمة بالكامل في المملكة العربية السعودية؛ لتلبية احتياجات الذكاء الاصطناعي التوليدي الوطنية، وتحقيق الريادة العالمية في تقنيات اللغة العربية، وبناء النماذج وتدريبها على بيانات شاملة؛ لخدمة احتياجات اللغة العربية والثقافة المحلية، وتوفير بنية تحتية سيادية متينة؛ لتطوير هذه النماذج محليًا، وتحقيق احتياجات السيادة التقنية.

ومن أبرز جهود سدايا في دعم اللغة العربية إطلاقها بالتعاون مع مجمع الملك سلمان العالمي للغة العربية مؤشر "بلسم" لتقييم وقياس نماذج اللغة العربية في مجالات البيانات والذكاء الاصطناعي، وذلك ضمن أعمال القمة العالمية للذكاء الاصطناعي في نسختها الثالثة التي عقدت خلال الفترة من 10 – 12 سبتمبر 2024 بالرياض.

ويأتي إطلاق المؤشر ضمن جهود "سدايا" وشركائها الإستراتيجيين لتطوير نماذج اللغة العربية التي تُعد مجالاً سريع التطور، ويسهم في تقييم النماذج الجديدة التي يتم العمل عليها باستمرار لإضافة مميزات وقدرات متقدمة.

وكان مجمع الملك سلمان للغة العربية قد عمل على تصميم إطار عمل متفق عليه لقياس مستوى نضج هذه النماذج في اللغة العربية، بدأ بوضع تصور للفكرة في عام 2022 قبل أن تتحول إلى مشروع بدأ العمل الفعلي عليه منذ العام الماضي 2023، واستهل المجمع تنفيذ المشروع بجمع البيانات المستهدفة، والعمل على تطوير النسخة الأولى لمؤشر القياس، "بلسم" بالتعاون مع "سدايا" وعدد من الجهات المشاركة وذات العلاقة في القطاعين الحكومي والخاص.

وأسهمت "سدايا" في إعداد معجم البيانات والذكاء الاصطناعي بالتعاون مع مجمع الملك سلمان للغة العربية لجمع أهم المصطلحات التقنية المتعلقة بالبيانات والذكاء الاصطناعي باللغتين العربية والإنجليزية من أجل زيادة المعرفة بتقنيات الذكاء الاصطناعي بكل أبعادها العالمية.

علاوة على ذلك، طورت "سدايا" نظام "صوتك"، الذي يعتمد على تقنيات الذكاء الاصطناعي لتحويل الكلام إلى نصوص باللغة العربية الفصحى واللهجات المحلية، مما يسهم في أتمتة تدوين الاجتماعات وتطوير أدوات محادثة ذكية. كما تم بناء أنظمة صوتية تفاعلية تُسهل التواصل وتعزز تجربة المستخدم.

ومن بين عدد من منتجات الذكاء الاصطناعي التي أطلقتها "سدايا" يبرز نموذج (علاّم) بوصفة أحد أبرز النماذج اللغوية للذكاء الاصطناعي التوليدي الذي يرُكّز على اللغة العربية وتم تطوير نموذج "علاّم" في مختبرات المركز الوطني للذكاء الاصطناعي التابع لسدايا بكوادر تقنية وطنية، مع الاستفادة من خبرات أكثر من 400 متخصص في مختلف المجالات، وبالتكامل مع أكثر من 160 جهة حكومية شاركت في بناء أكبر مجموعة بيانات عربية في العالم شملت أكثر من

500 مليار وحدة نصية عربية لتطوير وتدريب "علاّم".

وفي السياق ذاته نشرت سدايا مجموعة من التقارير والدراسات المعرفية التي تثري المحتوى العربي بمفاهيم الذكاء الاصطناعي وتطبيقاته، بهدف رفع مستوى الوعي المجتمعي بأهمية هذه التقنيات ودورها في تطوير القطاعات المختلفة.

وتأتي جهود "سدايا" في استخدام تقنيات الذكاء الاصطناعي لخدمة اللغة العربية لتؤكد التزامها بدعم اللغة العربية وتعزيز مكانتها تحقيقاً لعدد من الأهداف الطموحة للمحافظة على هذه اللغة، ودعمها ونشرها علومها في العالم.