من البيانات المفتوحة إلى البيانات الجاهزة للاستخدام عبر تقنيات وأنظمة الذكاء الاصطناعي

وضع الأسس من أجل استخدام الذكاء الاصطناعي على نحو مسؤول في مجال التنمية[1]

هايشان فو

رئيسة الخبراء الإحصائيين بالبنك الدولي ومديرة مجموعة بيانات التنمية- البنك الدولي

كريغ هامر

مدير برنامج أول لمجموعة بيانات التنمية في البنك الدولي

شهد إنتاج واستخدام بيانات التنمية تحولات جذرية خلال العقدين الماضيين. فقد أدى الانتقال من السجلات الورقية إلى التنسيقات الرقمية إلى تسهيل الوصول إلى البيانات وتبادلها أكثر من ذي قبل. ومع ظهور حركة البيانات المفتوحة، أصبح هناك توافر أكبر بكثير لمجموعات البيانات الحكومية والمؤسسية، وهو ما عزز فرص التحليل، ورفع مستوى الشفافية، ودعم الابتكار في مختلف المجالات. وفي الوقت نفسه، أسهمت التطورات الكبيرة في مجال البيانات الضخمة وعلوم البيانات في زيادة حجم المعلومات المتاحة وتنوعها، مما أتاح لصناع السياسات الاستفادة من رؤى أعمق وأكثر دقة لتوجيه سياسات التنمية.

واليوم، ومع التقدم السريع في تقنيات الذكاء الاصطناعي، وصلت بيانات التنمية إلى نقطة تحول جديدة: تطور بيانات التنمية وجاهزيتها لتقنيات الذكاء الاصطناعي. هذا التحول يعني أن بيانات التنمية أصبحت أكثر سهولة في الاكتشاف والفهم والوصول والاستخدام، سواء من قبل البشر أو تطبيقات الذكاء الاصطناعي.

لماذا نحتاج إلى بيانات جاهزة للاستخدام عبر أنظمة وتقنيات الذكاء الاصطناعي؟

يشهد الذكاء الاصطناعي، وخاصة نماذج اللغات الكبيرة (LLMs)، ثورة في الطريقة التي يتفاعل بها الناس مع البيانات. فاليوم، أصبح بإمكان جميع مستخدمي البيانات، بغض النظر عن مستوى خبرتهم – من المبتدئين إلى المحترفين – طرح أسئلة معقدة بلغة طبيعية على روبوتات الدردشة، مع توقع الحصول فوراً على رؤى دقيقة مستندة إلى البيانات، بالإضافة إلى تفسيرات واضحة وإجابات دقيقة.

[1] مدونة البنك الدولي، 07/21/2025، رابط.

حتى يتحقق هذا التحول والتطور، يجب أن تعمل أنظمة الذكاء الاصطناعي بكفاءة وموثوقية. وهذا يتطلب أن تمر البيانات التي تعتمد عليها هذه الأنظمة بعدة مراحل أساسية: التقييم، والتفسير، والتحقق من الصحة، والتنظيم، وأخيراً التبادل بطرق تدعم الاستخدام المسؤول والفعال للذكاء الاصطناعي. باختصار، لا بد أن تكون البيانات "جاهزة للاستخدام عبر أنظمة وتقنيات الذكاء الاصطناعي".

البيانات الجاهزة للاستخدام عبر أنظمة وتقنيات الذكاء الاصطناعي لا تلغي ما سبق اعتماده من تطورات أو مفاهيم أساسية أو معايير، مثل المبادئ الأساسية للإحصاءات الرسمية، أو أطر البيانات المفتوحة، أو مبادئ FAIR (إمكانية العثور على البيانات والوصول إليها وقابليتها للتشغيل البيني وإعادة استخدامها)، بل تبني عليها وتستفيد منها. فعندما نرسخ هذه الأسس والمعايير، تصبح البيانات الجاهزة للاستخدام من خلال أنظمة وتقنيات الذكاء الاصطناعي بيانات متاحة باستمرار لأغراض التنمية، وبالتالي ستكون بيانات مفتوحة ويسهل اكتشافها وإعادة استخدامها. كما يتم تنظيمها وتوثيقها بشكل منهجي، مما يضمن سهولة استخدامها من قبل الأفراد وأنظمة الذكاء الاصطناعي. لذلك، فإن ضمان جاهزية البيانات للاستخدام عبر تقنيات وأنظمة الذكاء الاصطناعي يمكن أن يختصر الطريق نحو اتخاذ قرارات أفضل بشأن بيانات التنمية، فضلاً عن المساهمة في تحسين السياسات وتسريع الابتكار، مما يجعل رؤى التنمية متاحة للجميع بشكل أكثر ديمقراطية. وفي هذا السياق، يواصل البنك الدولي جهوده ليصبح "بنك البيانات" الأكبر والأفضل، من خلال العمل المشترك مع البلدان الشريكة ومجتمع التنمية العالمي لتحقيق هذا الهدف.

أسباب تجهيز البيانات للاستخدام عبر تقنيات وأنظمة الذكاء الاصطناعي

أصبح الذكاء الاصطناعي التوليدي اليوم واجهة رئيسية يلجأ إليها الأفراد للبحث عن المعلومات، بما في ذلك ما يتعلق بالتنمية. وهناك منصات مثل Google's AI Overviews، وMicrosoft's Bing Perplexity.AI، وOpenAI's ChatGPT تقوم بالبحث عبر الإنترنت لتجميع مصادر مختلفة للمعلومات بهدف الإجابة الفورية عن استفسارات المستخدمين. ويكمن التحدي الحقيقي في أن إجابات الذكاء الاصطناعي لا يمكن أن تكون موثوقة إلا بقدر موثوقية البيانات التي تعتمد عليها. ففي كثير من الأحيان، تستند هذه الأنظمة إلى محتوى الإنترنت العام، بما في ذلك

المصادر غير الموثوقة أو نتائج البحث العشوائية، بدلاً من التركيز على مصادر البيانات الموثوقة مثل البنك الدولي أو مكاتب الإحصاء الوطنية.

ونظراً لأن أنظمة الذكاء الاصطناعي الحالية غالباً ما تختار مصادر بيانات ليست بالمستوى المطلوب، يواجه المستخدمون بشكل متكرر إجابات أصبحت متقادمة أو غير دقيقة، حتى عندما تكون هناك معلومات حديثة وصحيحة متاحة. وتكمن المشكلة في أن معظم إجابات الذكاء الاصطناعي تبدو وكأنها تقدم معلومات موثوقة، حتى عندما تكون في الواقع مجرد "هلوسات" أو معلومات غير صحيحة.

من المهم التأكيد على أن بيانات التنمية ذات الجودة العالية والموثوقة متوفرة وليست نادرة. بعبارة أخرى، لا ينبغي أن تتاح لأدوات الذكاء الاصطناعي مصادر بيانات ضعيفة أو غير موثوقة للإجابة عن الاستفسارات المتعلقة بموضوعات التنمية. وما نفتقده فعلاً هو وجود إطار معياري موحد وبنية تحتية قوية تتيح لأدوات الذكاء الاصطناعي الوصول بسهولة وبشكل مستمر إلى بيانات تنمية موثوقة من مصادر معتمدة، واستخدامها لتقديم إجابات دقيقة وموثوقة للمستخدمين.

إن توفير بيانات تنمية مجهزة للاستخدام من خلال تقنيات وأنظمة الذكاء الاصطناعي يمكن أن يسهم بشكل كبير في معالجة مشكلة موثوقية المعلومات وصحتها. ولتحقيق ذلك، يمكن دعم تقنيات وأنظمة الذكاء الاصطناعي بهدف الوصول بسهولة إلى بيانات التنمية الموثوقة من خلال تبني الحكومات والمنظمات الدولية والقطاع الخاص لبروتوكولات ومعايير تضمن قابلية التشغيل البيني بين الأنظمة المختلفة. هذا النهج من شأنه أن يدعم عملية اتخاذ القرار المبنية على الأدلة، ويعزز وصول الجمهور إلى معلومات موثوقة، كما يمكن أن يرسخ الثقة في المصادر المعتمدة لبيانات وإحصاءات التنمية.

ما الذي يجعل البيانات "جاهزة للذكاء الاصطناعي"؟

تكون بيانات التنمية جاهزة للذكاء الاصطناعي عندما يتم تنظيمها بشكل منهجي وتوثيقها بدقة، بحيث يكون معناها وسياقها واضحين ليس فقط للخبراء والمتخصصين، بل أيضاً للمستخدمين العاديين وأنظمة الذكاء الاصطناعي.

ثلاث ركائز أساسية تحدد بيانات التنمية الجاهزة للذكاء الاصطناعي:

١. أنظمة البيانات الجاهزة للذكاء الاصطناعي: تضمن البنية التحتية الأساسية، التي تشمل منصات الاكتشاف وواجهات برمجة التطبيقات والمعايير الفنية، ليس فقط تخزين البيانات، بل أيضاً سهولة اكتشافها، وقابليتها للتشغيل البيني، وإمكانية الوصول إليها.

٢. البيانات المعتادة والبيانات الوصفية ذات الجودة العالية: بيانات موثوقة، ومحدثة، وموثّقة بدقة، مصحوبة ببيانات وصفية شاملة ومنظمة. عندما يتعلق الأمر بتطبيقات الذكاء الاصطناعي، فإن ذلك يتطلب وجود مجموعات بيانات منظمة بشكل منهجي وموصوفة بتفاصيل كافية، لضمان قدرة المحللين – عنصر بشري أو آلة – على تفسير المعلومات بدقة ووضوح.

٣. الحوكمة القوية والشراكات الإستراتيجية: يُعد تطبيق سياسات شاملة، وإجراءات موحدة، وتعزيز التعاون بين مختلف القطاعات أمراً أساسياً لضمان سلامة البيانات، ودعم الشفافية، وتعزيز الاستخدام المسؤول لها. فهذه الخطوات ضرورية لبناء ثقة الجمهور في المعلومات المقدمة من البشر أو من خلال تقنيات وأنظمة الذكاء الاصطناعي.

عند الاستفادة من هذه العناصر الأساسية، تصبح بيانات التنمية مورداً متاحاً لجميع أصحاب المصلحة والأطراف المعنية وكل ذي شأن. فعندما تكون البيانات مهيأة وجاهزة للاستخدام عبر تقنيات وأنظمة الذكاء الاصطناعي، فإنها تعزز وصول الجمهور إلى المعلومات، وتوفر رؤى متقدمة بفضل تقنيات الذكاء الاصطناعي، كما تسهم في تسريع عملية اتخاذ القرار على نحو صحيح على مستوى المجتمع بأسره.

تحويل البيانات الجاهزة للاستخدام عبر تقنيات الذكاء الاصطناعي إلى واقع ملموس

يتطلب ذلك منا ترجمة المبادئ الأساسية إلى خطوات عملية قابلة للتنفيذ. وتشمل بيانات التنمية أشكالاً متعددة، مثل المؤشرات، والبيانات الجزئية، ومجموعات البيانات الجغرافية. ورغم أن التوصيات التالية يمكن تكييفها لتناسب مختلف أنواع البيانات، فقد صُممت بشكل خاص لتلائم المؤشرات ذات الصلة.

١. أنظمة البيانات الجاهزة للذكاء الاصطناعي

- **اكتشاف البيانات**: يتيح دمج إمكانات البحث الدلالي والمعجمي للمستخدمين وأنظمة الذكاء الاصطناعي تحديد البيانات ذات الصلة بسهولة، سواء استناداً إلى المعاني أو الكلمات الرئيسية.

كما أن دعم البحث متعدد اللغات يضمن وصول الجميع إلى المعلومات، مع توفير النتائج في تنسيقات مقروءة آلياً عبر واجهات برمجة التطبيقات.

- **الوصول إلى البيانات**: يعد تنفيذ المعايير المفتوحة والقابلة للتنفيذ آلياً، مثل معيار SDMX، خطوة أساسية لتمكين أنظمة الذكاء الاصطناعي من تفسير البيانات ودمجها بكفاءة. كما أن ضمان إتاحة البيانات بموجب تراخيص البيانات المفتوحة المسموح بها يعزز من سهولة الوصول إليها ويشجع على استخدامها على نطاق أوسع.
- **التشغيل البيني للذكاء الاصطناعي**: استخدام معايير مفتوحة، مثل بروتوكول (MCP)، لتمكين أنظمة الذكاء الاصطناعي من تحديد مصادر البيانات الموثوقة وتفسيرها بكفاءة، مع ضمان الشفافية ومواصلة الرقابة بشأن سياق البيانات واستخدامها.

تقوم مجموعة بيانات التنمية ومكتب كبير الخبراء الإحصائيين بالبنك الدولي باستثمارات نشطة في هذه المجالات، بما في ذلك تجربة أدوات بحث متقدمة، وتطوير نماذج تضمين لأغراض السياقات منخفضة الموارد، ودمج واجهات برمجة التطبيقات، وتطوير خادم لبروتوكول MCP لدعم منصة Data360 الجديدة ومجموعات البيانات المختارة الأخرى.

٢. البيانات المعتادة والبيانات الوصفية ذات الجودة العالية

- **ضمان جودة البيانات**: إجراء التحقق الشامل من صحة البيانات خلال جميع المراحل، باستخدام عمليات التحقق الآلية إلى جانب منهجيات الكشف عن البيانات الشاذة، مع ضمان حصول الموظفين على تدريب شامل على إدارة جودة البيانات، حيث إن التأكيد على قوة البيانات غاية في الأهمية لأغراض التحليل البشري أو باستخدام تقنيات الذكاء الاصطناعي.
- **تنسيقات متعددة للنشر**: عرض البيانات في مجموعة من التنسيقات المفتوحة، بما في ذلك CSV و Parquet و Arrow و JSON و APIs، لاستيعاب متطلبات المستخدم المتنوعة وتسهيل الدمج في تسلسل سير عمل تقنيات الذكاء الاصطناعي على نحو سلس.
- **استخدام معايير البيانات الوصفية**: تطبيق المعايير الدولية للبيانات الوصفية ومواصلة تحديث جميع البيانات الوصفية لمجموعات البيانات وعرضها بالتفصيل.

- **وضع معايير قوية للبيانات الوصفية**: صياغة وتنفيذ مبادئ توجيهية شاملة وعملية لتوليد البيانات الوصفية المنظمة، باستخدام الأدوات المستندة إلى الذكاء الاصطناعي لإجراء عمليات آلية تستهدف ضمان الجودة والتحسين.
- **أدوات الإدارة**: الاستثمار في المنصات والتقنيات المتقدمة التي تمكن من التحقق من صحة البيانات والبيانات الوصفية، فضلا عن الإدارة القوية للبيانات والبيانات الوصفية على نطاق واسع، من خلال الاستفادة من حلول الذكاء الاصطناعي.

ويعمل البنك الدولي، من خلال برامجه العملية الخاصة بجودة البيانات وبرامج الذكاء الاصطناعي لأغراض البيانات والبيانات لأغراض الذكاء الاصطناعي، على النهوض بهذه المبادرات من خلال توفير موارد مفتوحة المصدر، بما في ذلك محرر البيانات الوصفية، والمبادئ التوجيهية الشاملة لإنشاء بيانات وصفية عالية الجودة، وأطر تجريبية تستفيد من الذكاء الاصطناعي لتقييم جودة البيانات الوصفية وتعزيزها بكفاءة.

٣. الحوكمة والشراكات

- **الامتثال للسياسات والمساءلة**: تنفيذ سياسات قوية تشجع المعايير الصارمة لجودة البيانات والبيانات الوصفية والشفافية والوصول المفتوح للبيانات. ويشمل ذلك المتابعة المنتظمة للاستخدام من قبل الغير وإنشاء آليات فعالة لاستقاء الآراء التقييمية.
- **الأخلاقيات والخصوصية**: دمج الاعتبارات الأخلاقية وضمانات الخصوصية في كل مرحلة من مراحل التعامل مع البيانات، وإجراء تقييمات شاملة للآثار، وضمان الشفافية فيما يتعلق بالمنهجيات التحليلية ومصادر البيانات.
- **تضافر الجهود الدولية**: تسهيل المواءمة بين المعايير والمصطلحات من خلال مبادرات منسقة، وتعزيز عمليات المساعدة الفنية، وتطوير أدوات وموارد شاملة ومشتركة.
- **التعاون مع القطاع الخاص**: تعزيز الشراكات التعاونية مع شركات التكنولوجيا لدعم تطوير أدوات الذكاء الاصطناعي التي تعتمد على بيانات موثوقة ومنظمة بشكل جيد.

فمن خلال إطلاق مشروعات تجريبية مشتركة، ونشر أفضل الممارسات المتبعة، والدعوة إلى المزيد من الشفافية في جميع المبادرات، يمكن تحقيق تقدم ملموس في هذا المجال.

علاوة على ذلك، فإن دعم تطوير حلول الذكاء الاصطناعي منخفضة الموارد يضمن وصول هذه التقنيات إلى المؤسسات التي تعاني من قيود كبيرة في الموارد.

يعمل البنك الدولي حالياً على إقامة شراكات مع المنظمات الدولية، بما في ذلك اللجنة الإحصائية للأمم المتحدة، وصندوق النقد الدولي، ومنظمة التعاون والتنمية الاقتصادية، والبنك الأفريقي للتنمية، وعدد من البلدان، والقطاع الخاص لتعزيز الحوكمة واعتماد معايير وآليات عالمية لإدارة بيانات التنمية واستخدامها بشكل فعال مع أنظمة الذكاء الاصطناعي.

لماذا تتسم جاهزية الذكاء الاصطناعي لتوليد بيانات التنمية بالتفرد؟

تتميز بيانات التنمية عن معظم بيانات القطاع الخاص بأنها تخدم مجموعة واسعة من المستخدمين، مثل الحكومات، والمنظمات، والباحثين، والمجتمع المدني، وقطاع الأعمال، والجمهور العام.

وعندما تُعتبر هذه البيانات ذات منفعة عامة، يصبح من الضروري أن تتسم بالانفتاح والشفافية والمساءلة. ونظراً لأن بيانات التنمية تؤدي دوراً محورياً في توجيه السياسات واتخاذ قرارات الاستثمار في مختلف البلدان والأنظمة، فإن ضمان قابلية التشغيل البيني وتوفير توثيق شامل لها يعدان من الأمور الأساسية لتحقيق أقصى استفادة منها.

يحقق الاستخدام المستمر لبيانات التنمية وإعادة استخدامها قيمة مضاعفة. فعندما نقوم بتجهيز بيانات التنمية لاستخدامها عبر تقنيات وأنظمة الذكاء الاصطناعي وتوفير الحلول المدعومة بالذكاء الاصطناعي لمشكلات التنمية في القطاعين الحكومي والخاص، فإننا نعزز من تأثير هذه البيانات، ونشجع على تقاسم المنافع على نحو أكثر عدلاً وإنصافاً، كما نعزز الثقة في استخدام البيانات على نحو مسؤول. ويمكن للذكاء الاصطناعي أن يساعدنا في تحقيق قيمة اقتصادية واجتماعية أكبر من خلال الاستفادة من البيانات، وهو ما يطلق العنان لتحولات كبيرة في مختلف المجالات. وهذا بدوره يدعم جهودنا لتحسين مستويات المعيشة، ودفع عجلة التنمية الاقتصادية، والمساهمة في القضاء على الفقر.

دعوة للعمل

يُعد التحول نحو بيانات التنمية الجاهزة للذكاء الاصطناعي خطوة ضرورية وملحة على نطاق واسع. ويستلزم تحقيق هذا الهدف:

- الاستثمار في تطوير البنية التحتية للبيانات، وتنمية المهارات، وتبني المعايير العالمية المتعلقة بأنظمة البيانات، والبيانات الوصفية، والحوكمة.
- تعزيز التعاون بين الحكومات والمنظمات الدولية والقطاع الخاص، بهدف تيسير تبادل أفضل الممارسات وضمان الحفاظ على التنسيق الإستراتيجي بين جميع الأطراف المعنية.
- مواصلة الابتكار والمرونة نظراً للطبيعة المتطورة لتقنيات الذكاء الاصطناعي ومتطلبات المستخدمين.

ندعو المكاتب الإحصائية الوطنية، ومنتجي البيانات، وصناع السياسات، وشركاء التكنولوجيا إلى الانضمام والمشاركة الفاعلة في هذه المبادرة. فمن خلال التعاون المشترك والالتزام بالمعايير العالمية لجودة البيانات، نستطيع أن نضمن بقاء بيانات التنمية مورداً موثوقاً وشاملاً وقوياً يخدم الصالح العام، خاصة ونحن نمضي قدماً في عصر الذكاء الاصطناعي.

لنعمل معاً على تجهيز بيانات التنمية من أجل المستقبل، وضمان استفادة الجميع من فوائدها.