

البيانات المفتوحة: أساس للمنافع الرقمية والذكاء الاصطناعي النافع للجميع

د. كيبيل غرين، فريق Creative Commons



المخلص:

يناقش المقال أهمية البيانات المفتوحة للمنافع الرقمية العامة وتطوير الذكاء الاصطناعي الذي يخدم المجتمع، مع مراعاة الحالات التي تمنع فيها الخصوصية أو الأخلاق أو القانون مشاركة البيانات. ويعرض دور تحالف المنافع الرقمية العامة ومعايير قبول مجموعات البيانات والمحتوى، ومنها وضوح الترخيص وشموله، وسهولة الاكتشاف والوصول، وعدم التمييز في القيود التقنية. كما يناقش الجدل حول انفتاح بيانات تدريب النماذج، والتمييز بين نشر البيانات نفسها وتقديم وصف يتيح فهمها وإعادة الإنتاج. ويؤكد أن صعوبة الإتاحة لا تبرر التخلي عن تشجيعها حيث تكون ملائمة. وي طرح بحث نهج متدرج لمكونات الذكاء الاصطناعي، مع الحفاظ على معنى الانفتاح ودعم الاستقلال عن المنصات والابتكار والمنفعة العامة وعدم الإضرار.

الكلمات المفتاحية: البيانات المفتوحة، المنافع الرقمية العامة، الذكاء الاصطناعي المفتوح، التراخيص المفتوحة، التنمية المستدامة

تقر (Creative Commons) بأن بعض البيانات ينبغي ألا تُشارك لأسباب أخلاقية أو تتعلق بالخصوصية، وأن بعضها لا يمكن مشاركته لأسباب قانونية أو غيرها. ومع ذلك، ترى قيمة في تحفيز إنشاء البيانات المفتوحة وتبادلها واستخدامها لتعزيز إنتاج المعرفة.

ومع استمرار المجتمعات المفتوحة في تصور منافع رقمية وخدمات بنية عامة للتعليم والعلم والثقافة وتصميمها وبناءها، ينبغي، حيثما كان ممكناً وملائماً، أن تنتج بيانات مفتوحة أو تشاركها أو تبني عليها.

البيانات المفتوحة والمنافع الرقمية العامة:

(Creative Commons) عضو في تحالف (DPGA)، واعترف بأدواتها القانونية بوصفها منافع رقمية عامة (DPGs).

تشمل هذه المنافع برمجيات ومعايير وبيانات ونظم ذكاء ومجموعات محتوى مفتوحة، تراعي الخصوصية وأفضل الممارسات الأخرى، ولا تلحق الضرر، وترتبط بتحقيق أهداف التنمية المستدامة 2030. ولحل أكبر تحديات العالم، يحتاج الممولون والحكومات إلى الاستثمار فيها وتطويرها وترخيصها علناً ومشاركتها واستخدامها.

البيانات محرك للحياة الاقتصادية وقادرة على خدمة المنفعة العامة. ففي القطاع العام تفيد صنع السياسات وتقديم الخدمات وتوجيه الموارد المحدودة إلى الأشد حاجة، وتدعم مساءلة الحكومات والابتكار الاجتماعي؛ أي تستطيع تحسين حياة الناس. وعندما تُفلق أو تغيب، يفقد الجمهور هذه الفوائد.

شاركت المؤسسة حديثاً في لجنة فرعية للتحالف لصون المعنى الدقيق للانفتاح داخل معيار (DPG). وجاء تحديث ليقصر الاعتراف بمجموعات البيانات والمحتوى على ما يحمل تراخيص مفتوحة فعلياً، وفق شروط محددة.

أولها شمول الترخيص: يجب أن تخضع المجموعة كلها لترخيص مفتوح مقبول، ولا تُقبل المجموعات ذات التراخيص المختلطة.

وثانيها الإتاحة وقابلية الاكتشاف: تُتاح المجموعة المرخصة علناً بسهولة من موقع واحد محدد بوضوح، مثل عنوان (URL) فريد.

وثالثها القيود المسموح بها: يمكن اشتراط تسجيل الدخول أو التسجيل أو مفتاح (API) أو تقييد معدل الطلبات، بشرط ألا تميز هذه القيود بين المستخدمين أو تمنع الاستخدام بحسب الجغرافيا أو عوامل أخرى.

ويوضح التحالف أن ذلك يعزز الثقة، فيستخدم الناس الحلول دون مخاوف من التعدي على الحقوق، ويبسط الوصول والاستخدام والتبني الواسع، ويدعم ابتكاراً أقل غموضاً قانونياً.

الذكاء الاصطناعي والبيانات المفتوحة:

ترى المؤسسة أن للبيانات دوراً مماثلاً عند بحث إمكان كون الذكاء منفعة عامة تعالج تحديات عالمية. وتقدر جهود وضع التعريفات والتوصيات والإرشادات والتحذيرات في المجال سريع التطور.

بعد عامين من المشاورات، أطلقت (OSI) الإصدار 1.0 من تعريف الذكاء الاصطناعي المفتوح (OSAID) في 28 أكتوبر 2024، وهو خطوة مهمة لبدء الحوار. لكن متطلباته لمشاركة البيانات بقيت محل جدل، خاصة بشأن مشاركة بيانات التدريب وكيفيةها.

وترى (Creative Commons) أن صعوبة بناء مجموعات مفتوحة ونشرها لا تبرر التخلي عن تشجيعها. وعندما لا ينبغي نشر بيانات التدريب أو يتعذر ذلك، تدعو إلى أوصاف تفصيلية لمحتواها وتعليمات لإعادة الإنتاج، مع استمرار وصف البيانات نفسها بأنها مغلقة. أما ما يمكن فتحه ومشاركتها، فينبغي فتحه.

وتؤيد موقف ليف مارته نوردهوغ، الرئيسة التنفيذية للتحالف: يجب ألا يؤدي نهج أكثر تساهلاً مع نظم الذكاء إلى إضعاف حركة البيانات المفتوحة أو فتتها ضمن المنافع الرقمية. وقد يقلل رفع معيار انفتاح بيانات التدريب عدد النظم المؤهلة، لكن الارتباط بالتنمية والاستقلال عن المنصات ومراعاة عدم الإضرار منذ التصميم خصائص تميز هذه المنافع؛ ولهذا يلزم إدراج بيانات التدريب في الاعتبار.

الخطوات التالية:

ستواصل المؤسسة العمل مع التحالف والشركاء على معيار لتأهيل نماذج الذكاء بوصفها منفعة عامة. وستدافع عن مجموعات مفتوحة وعن بحث نهج متدرج يسمح باعتبار بعض مكونات النموذج منافع رقمية، دون اشتراط فتح كل مكون في النموذج الكامل.

والإرشادات المحدثة التي تقدّر نظم الذكاء المفتوحة بالكامل، القائمة على استخدام مجموعات مفتوحة ومشاركتها، جزء مهم من ضمان خدمة هذه النظم للمنفعة العامة.

المراجع:

1. معيار المنافع الرقمية العامة لدى (DPGA).
2. البنك الدولي، «بيانات من أجل حياة أفضل»، 2021، بترخيص (CC BY 3.0 IGO). وروابط المرجعين في صفحة المقال الأصلية.