الانفتاح وخصوصية البيانات: كيف نجمع بينهما؟

مايك بيرسل

الملخص:

يناقش المقال الجمع بين الانفتاح وخصوصية البيانات، مؤكدًا أن دعم المصدر المفتوح لا يعني إتاحة جميع المعلومات دون قيود. ويعرض أهمية حماية البيانات الشخصية وحدود استخدامها، وتأثير التنظيم الأوروبي في إجراءات المؤسسات، مع إبقاء توضيحات النطاق القانوني ضمن سياق المصدر. كما ينتقد الاستضافة التي تقايض التحكم بالبيانات بخدمات مجانية، ويسأل عن المستفيد الفعلي من استغلالها. ويستعرض البيانات الحكومية المفتوحة وتقنيات تسمح بتحليل مجموعات مع تقليل كشف الأفراد، ويميز بين البيانات العامة والمشتركة والسرية. ويقترح خطوات للتوعية والحماية والمراجعة وإعادة الاستخدام، مع الحفاظ على البيانات التي يجب أن تبقى خاصة. ويؤكد أن قرار الانفتاح يحتاج فهم الغرض والحقوق والمخاطر، لا التسليم بأن المجانية أو المشاركة مصلحة متطابقة لجميع الأطراف.

الكلمات المفتاحية: البيانات المفتوحة، خصوصية البيانات، المصدر المفتوح، حوكمة البيانات، التحكم بالمعلومات

اليوم هو يوم خصوصية البيانات، الذي يسمى في أوروبا «يوم حماية البيانات». وقد تظن أن العاملين في عالم المصدر المفتوح يرون وجوب إتاحة جميع البيانات بحرية، على أساس أن المعلومات تريد أن تكون حرة، لكن الحياة ليست بهذه البساطة، لسببين رئيسين:

أولًا، يعتقد معظمنا، داخل عالم المصدر المفتوح وخارجه، أن ثمة بيانات تخصنا لا نشعر بالارتياح إلى مشاركتها. وقد جمعت قائمة أمثلة في تدوينة سابقة.

ثانيًا، يعمل كثير منا في شركات تجارية أو مؤسسات تخضع لمتطلبات قانونية تحدد ما يمكنها مشاركته.

إذًا، خصوصية البيانات مهمة للجميع تقريبًا.

تختلف نقطة الانطلاق في تصور الناس والحكومات للبيانات التي ينبغي للمؤسسات استخدامها بين الولايات المتحدة وأوروبا. فالأولى تمنح عمومًا الكيانات، ولا سيما الشركات التجارية الكبيرة كما قد يقول المتشككون، حرية أوسع في استخدام ما جمعته من بيانات عنا كما تشاء. أما أوروبا فقد تبنت تاريخيًّا نظرة أشد تقييدًا؛ ويمكن القول إن نهجها سيصبح هو السائد في 25 مايو.

أثر اللائحة العامة لحماية البيانات:

هذا تعميم واسع، لكن الحقيقة أن ذلك اليوم هو موعد بدء إنفاذ تشريع أصدره الاتحاد الأوروبي سنة 2016، هو اللائحة العامة لحماية البيانات (GDPR). وتضع اللائحة قواعد صارمة لتخزين البيانات الشخصية، وأغراض استخدامها، ومن يطلع عليها، ومدة الاحتفاظ بها. كما تحدد معنى البيانات الشخصية، وهو نطاق واسع يمتد من اسمك وعنوان منزلك إلى سجلاتك الطبية وعنوان بروتوكول الإنترنت لحاسوبك.

والأهم أن (GDPR) لا تقتصر على الشركات الأوروبية؛ بل تشمل أي مؤسسة تعالج بيانات عن مواطني الاتحاد الأوروبي. فإذا كانت شركتك أرجنتينية أو يابانية أو أمريكية أو روسية وتجمع بيانات عن مواطن أوروبي، فإنها تخضع لها.

قد تقول: «أنا خارج الاتحاد، فما الذي يمكنهم فعله بي؟». الجواب بسيط: إن أردت الاستمرار في ممارسة الأعمال فيه فالأفضل الامتثال؛ لأن مخالفة القواعد قد تعرضك لغرامة تبلغ 4% من إيراداتك العالمية. نعم، الإيرادات العالمية، لا إيرادات دولة أوروبية بعينها، ولا الاتحاد وحده، ولا الأرباح. وهذه أرقام تستدعي الحديث مع الفريق القانوني، الذي سيحيلك إلى الإدارة التنفيذية، التي ستتجه سريعًا إلى فريق تقنية المعلومات لضمان الامتثال.

قد يبدو ذلك غير ذي صلة بغير مواطني الاتحاد، لكنه يعنيهم. ففي معظم الشركات سيكون من الأبسط والأكفأ تطبيق الحماية نفسها على بيانات جميع العملاء والشركاء والموظفين، بدل تخصيص إجراءات لمواطني الاتحاد وحدهم. وهذا أمر محمود بالتأكيد.

لكن تطبيق (GDPR) قريبًا على مؤسسات في أنحاء العالم لا يعني أن كل شيء أصبح على ما يرام. فنحن نمنح معلومات عن أنفسنا، وإذنًا باستخدامها، طوال الوقت.

هناك عبارة معبرة، وإن كانت محل خلاف: «إذا لم تدفع، فأنت المنتج». ومعناها أن طرفًا آخر قد يدفع لاستخدام بياناتك حين تستخدم الخدمة مجانًا. هل تدفع مقابل فيسبوك أو تويتر أو (Gmail)؟ كيف تجني هذه الخدمات المال؟ جزئيًّا من الإعلان، الذي قد يعده بعضهم خدمة لك، لكن الواقع أنهم يستخدمون بياناتك للحصول على أموال المعلنين. أنت لست عميلًا للإعلان نفسه؛ ولا تصبح عميلًا للمعلن إلا إذا اشتريت منه. وحتى ذلك الحين تكون العلاقة بين مالك منصة الإعلان والمعلن.

تسمح بعض الخدمات بالدفع لتقليل الإعلانات أو حذفها، مثل (Spotify). ومن جهة أخرى، قد تُضاف الإعلانات حتى إلى خدمات تظن أنك تدفع مقابلها؛ فأمازون، مثلًا، تعمل بحسب ما يتداول على إتاحة إعلانات عبر (Alexa). ما لم نكن مستعدين للدفع مقابل كل هذه الخدمات «المجانية»، فعلينا إدراك ما نتنازل عنه، واتخاذ قرارات بشأن ما نكشفه وما نحجبه.

من هو العميل:

هناك مسألة أخرى ينبغي أن تشغلنا، تنشأ مباشرة عن أحجام البيانات المنتجة. فمؤسسات كثيرة، ومنها مؤسسات «عامة» كالجامعات والمستشفيات والدوائر الحكومية، تنتج باستمرار كميات هائلة من البيانات ولا تملك قدرة كافية لتخزينها. ولو افتقرت إلى قيمة طويلة الأمد لكان الأمر مختلفًا، لكنها ذات قيمة؛ فأدوات معالجة البيانات الضخمة تتطور، والمؤسسات تدرك إمكان تنقيبها الآن وفي المستقبل.

ومشكلتها أن حجم البيانات يزيد أسرع من قدرة التخزين: ماذا تفعل بها؟ ولحسن الحظ ـ وأقول ذلك بقدر كبير من السخرية ـ تتقدم الشركات الكبرى للمساعدة: «أعطونا بياناتكم، وسنستضيفها مجانًا، بل سنسمح لكم باستخدام ما جمعتموه متى أردتم!» أليست صفقة رائعة؟ أليس مثالًا رائعًا على تبني الشركات الكبرى موقفًا خيريًّا لمساعدة المؤسسات العامة التي جمعت كل هذه البيانات الثمينة عنا؟

للأسف، ليس العمل الخيري الدافع الوحيد. فلهذه الاستضافة ثمن: تحصل الشركات، مقابلها، على حق بيع الوصول إلى البيانات لأطراف ثالثة. وهل تظن أن المؤسسات العامة أو الأشخاص الذين جُمعت بياناتهم سيكون لهم رأي في تحديد هوية هذه الأطراف أو كيفية استخدام البيانات؟ سأترك الجواب تمرينًا للقارئ.

الانفتاح والجوانب الإيجابية:

ليست الأخبار كلها سيئة. هناك حركة متنامية للبيانات المفتوحة بين الحكومات تشجع الإدارات على إتاحة جانب كبير من بياناتها مجانًا للجمهور والجهات الأخرى، وأحيانًا ينص التشريع صراحة على ذلك. وبدأت مؤسسات تطوعية كثيرة، وبخاصة التي تتلقى تمويلًا عامًا، في فعل الشيء نفسه، بل تظهر بوادر اهتمام لدى مؤسسات تجارية.

وتبرز تقنيات، منها الخصوصية التفاضلية والحوسبة متعددة الأطراف، بدأت تتيح تنقيب مجموعات بيانات معًا من دون كشف الكثير عن الأفراد؛ وهذه مهمة حاسوبية كانت تاريخيًّا أصعب مما قد تتوقع.

ماذا يعني ذلك لنا؟ كتبت من قبل في (Opensource.com) عن المشاع في المصدر المفتوح، وأزداد اقتناعًا بوجوب النظر إلى ما وراء البرمجيات، إلى العتاد والمؤسسات والبيانات، وهي موضوعنا هنا. لنتخيل شركة «أ» تقدم خدمة إلى شركة عميلة «ب». لدينا أربعة أنواع من البيانات:

1. بيانات مفتوحة تمامًا: تراها «أ» و«ب» والعالم كله.

2. بيانات معروفة ومشتركة وسرية: تراها «أ» و«ب» وحدهما.

3. بيانات سرية للشركة: تراها «أ» ولا تراها «ب».

4. بيانات سرية للعميل: تراها «ب» ولا تراها «أ».

ربما ينبغي أولًا أن نكون أكثر انفتاحًا، وأن نجعل الفئة الأولى هي الوضع الافتراضي. فقد تفيد البيانات عن السيارات الذاتية القيادة، والتعرف إلى الصوت، والرواسب المعدنية، والإحصاءات السكانية، فائدة هائلة لو أتيحت للجميع. ألن يكون رائعًا كذلك أن نجد طرقًا تنقل بعض بيانات الفئات الثانية والثالثة والرابعة إلى الفئة الأولى، مع إبقاء التفاصيل سرية؟ هذا ما تأمله بعض التقنيات الجديدة قيد البحث. وما زالت بعيدة بعض الشيء، فلا تندفع في التفاؤل، لكن ابدأ بالتفكير في جعل قدر أكبر من بياناتك مفتوحًا افتراضيًّا.

خطوات عملية:

ماذا نستطيع فعله للجمع بين الخصوصية والانفتاح؟ هذه خطوات خطرت لي، وأرجو إضافة المزيد في التعليقات:

• تحقق من جدية مؤسستك في التعامل مع (GDPR)؛ وإن لم تكن جادة فادفعها إلى ذلك.

• اجعل تشفير البيانات الحساسة هو الأصل، واستخدم التجزئة حين تلائم، واحذفها عندما تنتفي الحاجة. فلا عذر اليوم لبقاء البيانات مكشوفة إلا أثناء معالجتها فعلًا.

• فكر فيما تفصح عنه عند التسجيل في الخدمات، وبخاصة وسائل التواصل الاجتماعي.

• ناقش الأمر مع أصدقائك غير المتخصصين في التقنية.

• علّم أبناءك وأبناء أصدقائك وأصدقاءهم. والأفضل أن تتحدث إلى معلميهم وتقدم عرضًا في مدارسهم.

• شجع المؤسسات التي تعمل فيها أو تتطوع لها أو تتعامل معها على فتح بياناتها افتراضيًّا. ابدأ بالسؤال «لماذا لا أجعل هذه البيانات عامة؟» بدل «لماذا أجعلها عامة؟».

• جرب الوصول إلى بعض مصادر البيانات المفتوحة: نقّبها، أو اصنع تطبيقات تستخدمها، أو حللها إحصائيًّا، أو ارسم مخططات جميلة، أو اصنع موسيقى مثيرة للاهتمام. المهم أن تفعل بها شيئًا، ثم أخبر الجهات التي أتاحتها، واشكرها وشجعها على المزيد.

المراجع:

• https://en.wikipedia.org/wiki/Data_Privacy_Day

• https://en.wikipedia.org/wiki/Information_wants_to_be_free

• https://aliceevebob.wordpress.com/2017/06/06/helping-our-governments-differently/

• https://www.outpost9.com/reference/jargon/jargon_40.html#TAG2036