إدارة بيانات البحث: مسؤوليات المكتبات وحدود دورها

كيفن آشلي



الملخص:

يناقش المقال دوافع الإدارة الجيدة لبيانات البحث، من إعادة الاستخدام في البحث والتعليم والسياسات والأعمال إلى الامتثال والنزاهة وحماية السمعة وتعزيز الأثر العلمي. ويؤكد ضرورة الموازنة بين تكلفة جمع البيانات وحفظها وقيمتها المحتملة، بدل افتراض الاحتفاظ بكل شيء دائمًا. ويعرض أدوار المكتبات في التقييم والاختيار، ومسح الموجود وتحديد المسؤوليات، وتشغيل سجل مؤسسي للبيانات يتكامل مع المنح والمنشورات وبنى الاكتشاف. كما يوضح إمكان قيادتها التنسيق بين الباحثين وتقنية المعلومات ومكاتب البحث، والاستفادة من أدوات مركز الإشراف الرقمي. ويحذر من العمل المنفرد أو منافسة مراكز البيانات التخصصية القائمة، ويدعو إلى شراكات مؤسسية تجعل البيانات قابلة للحفظ والاستشهاد والاكتشاف وإعادة الاستخدام بما يخدم الباحثين ومؤسساتهم والناشرين.

الكلمات المفتاحية: إدارة بيانات البحث، المكتبات البحثية، إعادة استخدام البيانات، سجل البيانات المؤسسي، الإشراف الرقمي

لماذا تتزايد أهمية القضية:

شهدت السنوات الأخيرة اهتمامًا وتمويلًا كبيرين للإدارة الفعالة لبيانات البحث من الحكومات وجهات تمويل البحث والباحثين ومؤسساتهم وأطراف أخرى. فما الأدوار التي يمكن أن يتولاها المكتبيون والناشرون في هذا المجال؟ وهل تقع عليهم مسؤوليات أصلًا؟ ما زالت بعض جوانب هذا النقاش دون حسم. سأحاول أولًا تفسير أهمية القضية المتنامية، وخاصة للجامعات التي يجري فيها البحث، ثم أطرح رأيًا في الأدوار التي ينبغي أن تتولاها المكتبات، وتلك التي ينبغي تجنبها. وآمل أن يتضح أن كثيرًا من دوافع إدارة البيانات إدارة سليمة لا يختص ببيانات البحث وحدها.

لماذا نهتم:

الدافع الأول، الذي يكفي وحده للاهتمام بالمسألة، هو إعادة استخدام البيانات. فجمع البيانات مكلف ويمثل استثمارًا. وكثير منها قائم على الرصد ولا يمكن تكراره، سواء جاء من قمر صناعي يراقب طقس الأرض، أو استطلاع لنوايا التصويت في غلاسكو في أوائل مايو 2012، أو تقارير الجمهور عن الطيور في حدائقهم في يوم معين. في كل حالة توجد أداة يمكن استخدامها مرارًا، لكن البيانات الناتجة تختلف عادة. لذلك يمكن القول إن قيمتها تتجاوز تكلفة جمعها وحدها.

ولكل البيانات، رصدية كانت أم غير ذلك، إمكان لإعادة الاستخدام. ويتفاوت هذا الإمكان كما تتفاوت تكلفة الاحتفاظ، فلا يعني ذلك ضرورة حفظ كل شيء إلى الأبد. بل ينبغي اتخاذ قرارات مدروسة توازن تكلفة إنتاج البيانات وحفظها بالقيمة المحتملة لإعادة استخدامها. وفي بعض الحالات يكون التخلص منها وإنتاجها مجددًا عند الحاجة أقل تكلفة فعلًا.

وينبغي أيضًا النظر إلى إعادة الاستخدام خارج الإطار الضيق للتخصص الذي أنتجها؛ فكثير من الحالات الأشد إثارة للاهتمام يأتي من البحوث المشتركة بين التخصصات. ويمكن استخدام البيانات خارج البحث الأكاديمي أيضًا، كالتعليم والتعلم، وهما مهمان أساسًا للجامعات. وغالبًا ما يكون ترسيخ الممارسة الجيدة في جمع البيانات ووصفها واختيارها وتحليلها لدى طلاب اليوم أيسر بأمثلة حقيقية، جيدة ورديئة، جُمعت لأغراض بحثية، بدل أمثلة مصطنعة.

وقد تفيد البيانات خارج الجامعة في توجيه السياسات العامة ودعم الأنشطة التجارية؛ وبالمثل تكون بيانات هذه المجالات غير الأكاديمية ذات قيمة للبحث الأكاديمي. ويشبه ذلك بوضوح إعادة استخدام البيانات الإدارية والببليوجرافية وبيانات أنشطة المكتبات لأغراض لا تعود كلها بنفع مباشر على المكتبات نفسها.

الامتثال والنزاهة والأثر العلمي:

هناك أسباب أخرى لأهمية الإدارة الجيدة. فقد تفرض المتطلبات القانونية والتنظيمية إدارة سليمة للبيانات [1]؛ وفي المملكة المتحدة تشمل قوانين حماية البيانات وحرية المعلومات وشروط الجهات الممولة. فسيفرض مجلس بحوث العلوم الهندسية والفيزيائية (EPSRC)، وهو من أكبر جهات التمويل البريطانية، حفظ بيانات مشروعاته 10 سنوات بعد آخر استخدام لها؛ وقد يعني ذلك حفظ البيانات كثيرة الاستخدام إلى الأبد. وتوجد أيضًا متطلبات خاصة بمجالات مثل التجارب السريرية.

وتمثل سلامة ممارسة البحث وسمعة مؤسساته دافعين قويين. فإذا دعمت البيانات استنتاجات بحث منشور، فمن المهم أن تتاح لمن يريد مساءلة البحث أو التحقق منه. ويشترط كثير من الناشرين والتخصصات الآن أن تودع بيانات المنشور لدى جهة تكفل حفظها الدائم وإتاحتها للآخرين. ولو طُبق ذلك سابقًا لانكشفت حالات كثيرة من سوء السلوك البحثي في وقت أبكر.

وتذهب بعض الدوريات، مثل دوريات الاتحاد الدولي لعلم البلورات [2]، إلى أبعد من ذلك؛ إذ يكون الملف المقدم الواحد هو المقال المنشور والبيانات الداعمة له معًا، مما يتيح استكشاف البيانات مباشرة عبر المنشور الإلكتروني.

وأظهرت دراسات أن إتاحة البيانات الداعمة للبحث [3، 4]، وإقامة روابط فعالة بين المنشور والبيانات المحفوظة عادة في نظم مختلفة للإشراف على البيانات، تعززان أثر البحث. وينتج عن ذلك مزيد من الاستشهادات بالبحث والبيانات وإعادة استخدامهما. وهي فائدة للباحثين ومؤسساتهم والناشرين، ومن مصلحتهم جميعًا التعاون لجعل الاستشهاد بالبيانات أوسع انتشارًا وأيسر وأشد فعالية.

يمكن تلخيص الدوافع في فئتين: عواقب سلبية يسببها إهمال الإدارة الجيدة، ومنافع يحققها الالتزام بها.

ماذا نفعل:

إذا اقتنعتم بأهمية القضية، فآمل أن تسألوا عما تستطيعون أنتم أو مؤسساتكم فعله. وهذا أحد المجالات التي تهدف مؤسستي، مركز الإشراف الرقمي (DCC)، إلى المساعدة فيها. فنقدم إرشادات وأدوات لجميع جوانب دورة إدارة البيانات، ولتطوير القدرات والعمليات اللازمة داخل المؤسسات.

حددنا أدوارًا تكون المكتبات مؤهلة لها غالبًا. وبعضها يتصل بمهارات قائمة في مكتبات البحث، مثل التقييم والاختيار. فما دام لا يمكن حفظ كل شيء إلى الأبد، يجب اتخاذ قرارات مستنيرة بشأن ما يُحفظ ومدة حفظه. ومع أن بعض العوامل يحتاج إلى معرفة تخصصية لا يمتلكها إلا منشئو البيانات، فإن مبادئ عامة كثيرة تدخل في القرار، وتستطيع المكتبات الإسهام في تطويرها ضمن السياسة المؤسسية.

ومهمة أخرى هي المسح الأولي للبيانات الموجودة وتحديد المسؤول عنها. وقد طور (DCC) إطار تدقيق البيانات (DAF) للمساعدة على تنظيم ذلك. وتوضح الخبرة أن معظم المؤسسات، بل معظم الأقسام البحثية، لا تفهم جيدًا ما تملكه من بيانات، ومن المسؤول عنها الآن ومن ينبغي أن يتحمل مسؤوليتها مستقبلًا. وقد يعوق هذا الجهل الأساسي التخطيط للقدرات، ويسبب تبعات قانونية خطيرة حين ترد طلبات اطلاع تتعلق بحقوق الأفراد أو بحرية المعلومات.

وتتطلب إدارة البيانات على المدى الطويل سجلًا مؤسسيًا لها، يمكنه إظهار معلومات عن مخرجات البحث للآخرين، وأن يكون مقصدًا لروابط البيانات الواردة من المنشورات. وتشغيل هذا السجل مجال تستطيع المكتبات قيادته، بالتعاون الفعال مع مكاتب البحث التي ستزداد حاجتها إلى ربط مجموعات البيانات بالمنح التي مولت إنشاءها. وبصيانة سجلات تتكامل مع بنى اكتشاف البيانات الوطنية والدولية، تؤدي المكتبات دورًا رئيسًا في ضمان حفظ البيانات وقابليتها للاستشهاد بها والعثور عليها وإعادة استخدامها.

ويمكن للمكتبات أن تقود الجهود الرامية إلى تحفيز العمل المؤسسي والاستجابة للمتطلبات الخارجية، مثل متطلبات (EPSRC). لا تستطيع حل القضايا منفردة، لكنها تستطيع جمع خدمات تقنية المعلومات والحوسبة البحثية وإدارة البحث والباحثين لإيجاد حلول تشمل المؤسسة. ويعمل (DCC) مع جامعات عدة لهذا الغرض، وسيقدم دراسات حالة تساعد غيرها على تكرار الممارسة الجيدة. وإلى ذلك الحين، تساعد أداة (CARDIO) في تحديد نطاق المشكلة وما هو قائم والخطوات التالية المناسبة.

ما الذي ينبغي تجنبه:

توجد تخصصات بحثية لديها مراكز بيانات جيدة تؤدي المهام السابقة وغيرها على نحو ممتاز. وحيث تتوافر، يرغب الباحثون في استخدامها، فلا ينبغي إعاقتهم عن ذلك، بل يمكن مساعدتهم على تيسير العملية. وقد تظل المؤسسة بحاجة إلى سجل بيانات وصفية يحيل إلى نسخة مركز البيانات. وينبغي كذلك تجنب العمل منفردين، وخاصة دون دعم الباحثين من مستوى نائب المسؤول الأكاديمي فما دونه.

تذكروا أن أداء العمل جيدًا يفيدكم ويفيد مؤسستكم، وأن أداءه على نحو سيئ أو إهماله قد يعرض سمعتكم وتمويلكم المقبل للخطر. وتتوافر المواد المذكورة وغيرها على موقع مركز الإشراف الرقمي (DCC).

المراجع:

1. DCC Policy & Legal resources. http://www.dcc.ac.uk/resources/policy-and-legal (اطلاع: 28 مايو 2012).
2. Crystal Structure Communications Online. http://journals.iucr.org/c/ (اطلاع: 28 مايو 2012).
3. Piwowar, H., Day, R., & Fridsma, D. (2007). Sharing Detailed Research Data Is Associated with Increased Citation Rate. PLoS ONE, 2(3), e308. (اطلاع: 28 مايو 2012).
4. Pienta, A. M., Alter, G. C., & Lyle, J. A. The Enduring Value of Social Science Research: The Use and Reuse of Primary Research Data. http://hdl.handle.net/2027.42/78307 (اطلاع: 28 مايو 2012).