

بيانات البحث والمكتبات: من يفعل ماذا؟

كيفن آشلي

ورقة مبنية على عرض في مؤتمر UKSG الخامس والثلاثين، غلاسكو، مارس 2012. تعكس السياسات والأمثلة وقت نشر الأصل، وأعدت هذه الترجمة لاحقاً ضمن استكمال أرشيف المجلة.

الملخص:

تجعل ضغوط خارجية متعددة إدارة بيانات البحث RDM موضع اهتمام متزايد لدى القيادات العليا في الجامعات والمؤسسات البحثية. لكن، إلى جانب هذه الضغوط، توجد أسباب وجيهة لإنشاء خدمات فعالة لإدارة البيانات داخل المؤسسات، تعود بالفائدة على الباحثين ومؤسساتهم ومن ينشرون بحوثهم. تعرض المقالة بعض هذه الدوافع الإيجابية والسلبية، وتوضح كيف تؤدي المكتبات دوراً، بل قد تقود تطوير خدمات تعمل داخل المؤسسة وضمن بنية وطنية ودولية لبيانات البحث.

لماذا تتراد أهمية القضية؟:

شهدت السنوات الأخيرة اهتماماً وتمويلاً كبيرين للإدارة الفعالة لبيانات البحث من الحكومات وجهات تمويل البحث والباحثين ومؤسساتهم وأطراف أخرى. فما الأدوار التي يستطيع المكتبيون والناشرون القيام بها، إن كانت لهم أدوار؟ كان ذلك موضع نقاش لم يُحسم كله. سأحاول أولاً تفسير أهمية القضية المتنامية، وخاصة للجامعات التي يجري فيها البحث، ثم أطرح رأياً في الأدوار التي ينبغي أن تتولاها المكتبات، وتلك التي ينبغي تجنبها. وآمل أن يتضح أن كثيراً 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 استثماراً. 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 مراراً، لكن البيانات الناتجة تختلف عادة. لذلك يمكن القول إن قيمتها تتجاوز تكلفة جمعها البسيطة.

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 مجدداً عند الحاجة أقل تكلفة فعلاً.

وينبغي أيضاً 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 أيضاً، كالتعليم والتعلم، وهما مهمتان أساساً للجامعات. وغالباً ما يكون ترسيخ الممارسة الجيدة في جمع البيانات ووصفها واختيارها وتحليلها لدى طلاب اليوم أيسر بأمثلة حقيقية، جيدة وردية، 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. وبشبه ذلك بوضوح إعادة استخدام البيانات الإدارية والبيولوجرافية 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. فسيُفرض 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 مسألة البحث أو التحقق منه. ويشترط كثير من الناشرين والتخصصات الآن أن توضع بيانات المنشور في جهة تمنحها موطئًا دائمًا وتتيحها 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 رعاية مختلفة، تعززان أثر البحث. وينتج عن ذلك مزيد من الاستشهادات بالبحث والبيانات وإعادة استخدامها. وهي فائدة للباحثين ومؤسساتهم والناشرين، ومن مصلحتهم جميعًا التعاون لجعل الاستشهاد بالبيانات أوسع انتشارًا وأيسر وأشد فعالية.

يمكن تلخيص الدوافع في فئتين: أسباب تجعل أمورًا سيئة تقع إن أهملنا الإدارة الجيدة، وأخرى تجعل من ممارستها يجني أمورًا حسنة.

ماذا نفعل؟:

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

حددنا أدواتًا 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 مستقبلاً. وقد يعوق هذا الجهل الأساسي التخطيط للقدرات، ويسبب تبعات قانونية خطيرة حين ترد طلبات اطلاع تتعلق بحقوق الأفراد أو بحرية المعلومات.

وتحتاج المعالجة السليمة طويلة الأمد إلى سجل مؤسسي للبيانات، يمكنه إظهار معلومات عن مخرجات البحث للآخرين، وأن يكون مقصودًا 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 والعثور وإعادة الاستخدام.

ويمكن للمكتبات قيادة تحفيز العمل المؤسسي والاستجابة للضغوط 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 تتوفر، يرغب الباحثون في استخدامها، فلا ينبغي الوقوف في طريقهم، مع إمكان تبسيط العملية. وقد تطل 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 مؤسساتكم، وأن أداءه على نحو سيئ أو إهماله قد يعرض سمعتكم وتمويلكم المقبل للخطر. وتتوافر المواد المذكورة وغيرها على موقع DCC: www.dcc.ac.uk

مراجع الأصل:

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).

المؤلف:

كيفن آشلي، مدير مركز الإشراف الرقمي Digital Curation Centre وقت نشر الأصل.

البيانات البليوجرافية للمقال الأصلي:

Ashley, K. (2012). Research data and libraries: who does what. Insights, 25(2), 155-157. <https://doi.org/10.1629/2048-7754.25.2.155>

تاريخ نشر الأصل: 2012-07-05. ترجمة عربية أُعدت بمساعدة الذكاء الاصطناعي ضمن استكمال أرشيف المجلة. حقوق الأصل للمؤلفين المذكورين. ترخيص النص والترجمة: المشاع الإبداعي نسب المصنف (CC BY 4.0). حُفظ النص والمراجع؛ لم تُكرر الاقتباسات الجانبية المستقلة من المتن، وحُذفت صورة المؤلف الشخصية وبيانات الاتصال البريدية. أُصلح التنسيق المطبوعي لمرجع Piwowar دون تغيير مضمونه.

<https://creativecommons.org/licenses/by/4.0/>