

ثلاثة مسارات ووجهة واحدة: إدارة بيانات البحث ومبادئ (FAIR) والانفتاح

روزي هيغمان، دانيال بانفرت، سارة جونز



المخلص:

يتناول مقال الرأي هذا ثلاثة مفاهيم متداخلة ومتميزة في التعامل مع بيانات البحث: إدارة بيانات البحث، ومبادئ (FAIR) (قابلية البيانات للعثور عليها والوصول إليها والتشغيل البيئي وإعادة الاستخدام)، والبيانات المفتوحة. ويحدد المؤلفون حدود كل مفهوم ومواضع تقاطعه مع الآخرين، ويصححون مفاهيم خاطئة شائعة، منها أن بيانات (FAIR) يجب أن تكون مفتوحة، وأن البيانات المفتوحة أو (FAIR) جيدة الجودة بالضرورة. ويقترحون نموذجًا ثلاثي الأبعاد ينظر إلى كل مفهوم بوصفه طبقًا متدرجًا، ويناقشون حالتين توضيحتين: بيانات منشورة تفي جزئيًا بالمتطلبات لكنها غير قابلة للاستخدام، وبيانات نموذجية مغلقة إلى جانب شيفرة مفتوحة. ثم يقدمون خمس توصيات لأمناء المكتبات وأمناء البيانات في توعية الباحثين. ويخلصون إلى أن إدارة البيانات شرط سابق لا غنى عنه لتحقيق (FAIR) والانفتاح، وأن هذين المبدأين يمكن أن يحقّزا الباحثين على الممارسات الجيدة منذ البداية.

الكلمات المفتاحية: إدارة بيانات البحث، مبادئ (FAIR)، البيانات المفتوحة، العلم المفتوح، التوعية بإدارة البيانات

مقدمة:

شهدت السنوات العشرين الماضية عدة تحولات في مواضع التركيز والأولويات في مجال إدارة بيانات البحث ومشاركتها. فقد تطورت سياسات الجهات الممولة للبحث وتعززت على مر السنين، من تطلعات مبهمّة إلى متطلبات واجبة التنفيذ تصحبها أنشطة للامتثال والمتابعة. وعلى وجه الخصوص، حدث تحول في الخطاب من التركيز على إدارة بيانات البحث (ولا سيما من خلال خطط إدارة البيانات) إلى إدراج مشاركة البيانات والوصول إليها. ولو توسّعنا في تحليل هذا التحول، لتبيّن أن السياسات الحديثة لا تشترط أن تكون البيانات مُدارة ومفتوحة (حيثما أمكن) فحسب، بل أن تكون أيضًا متوافقة مع مبادئ (FAIR) (قابلة للعثور عليها، وقابلة للوصول إليها، وقابلة للتشغيل البيئي، وقابلة لإعادة الاستخدام).

قلّما حظي شيء في مجال بيانات البحث بما حظيت به مبادئ (FAIR) للبيانات من رواج وقبول عام؛ وقد نشأت هذه المبادئ في مؤتمر لورنتز (Lorentz) في يناير 2014، ثم طُرحت للتشاور ونُشرت أول مرة تحت مظلة (FORCE11). [1] ومع تفاوت التفسيرات لمعنى أن يكون الشيء متوافقًا مع (FAIR) أو لمدى توافق كائن ما معها، فلا أحد يختلف مع ما تقرره المبادئ. فمبادئ (FAIR) تجمع في الواقع أفكارًا لها تاريخ طويل في مبادئ منظمة التعاون الاقتصادي والتنمية (OECD) [2] وبيان وزراء العلوم في مجموعة الثماني (G8)، [3] وتضم هذه العناصر معًا بطريقة موجزة وواضحة، تحت اختصار جذاب. ويمكنها أن تفتح حوارات مع الباحثين والجهات الممولة بطرق لم يفلح فيها قط الخطاب التقليدي عن إدارة البيانات، الذي قد يبدو مملًا.

إن إدارة بيانات البحث ومبادئ (FAIR) والانفتاح ثلاثة مفاهيم متداخلة ولكنها متميزة. ويأتي كل منها بتركيز ونقطة قوة مختلفين، وثمة مجال واسع للإثراء إذا طبقت مجتمعة. فإدارة بيانات البحث هي الأساس؛ إذا لم تُنشأ البيانات وتُدر على النحو السليم في المراحل الأولى من البحث، فسيكون من الصعب جدًّا جعلها متوافقة مع (FAIR) أو مفتوحة. وستؤثر ملكية البيانات وتوثيقها والصيغ والمعايير المستخدمة جميعًا في القدرة على المشاركة الفعالة، وكثيرًا ما تُحدد هذه الخيارات قبل إتاحة المخرجات النهائية بوقت طويل.

تمكّن إدارة البيانات من المشاركة وفق مبادئ (FAIR) والانفتاح، في حين يمكن لمبادئ (FAIR) والانفتاح أن تكون مصدر إلهام لإشراك الباحثين في الإدارة الفعالة للبيانات. وكثيرًا ما يرغب الباحثون في أن يكونوا متوافقين مع (FAIR)، وأحيانًا في أن يكونوا منفتحين؛ فهذان تطلعان نبيلان. أما إدارة البيانات فهي على العكس أشبه بـ«البطة القبيحة»، إذ يُنظر إليها على أنها عمل شاق قليل الجاذبية يعرف الناس أنه ينبغي لهم القيام به لكنهم لا يرغبون كثيرًا في الانخراط فيه. [4] وباستخدام لغة (FAIR) والانفتاح الأكثر جاذبية، يمكننا أن نشرك الناس في إدارة البيانات أيضًا.

إدارة بيانات البحث ومبادئ (FAIR) والانفتاح:

مع إدخال مفاهيم مثل (FAIR)، تبرز الحاجة إلى معالجة العلاقة بينها وبين الأفكار الراسخة الأخرى. ويمكن أن يساعد توفير مزيد من الوضوح حول تقاطعات إدارة بيانات البحث ومبادئ (FAIR) والانفتاح على إدراك مواضع التوافق وتحديد الثغرات في الوعي والدعم. وسيستعرض هذا القسم بإيجاز كلاً من المفاهيم الثلاثة، ويقترح طرقًا لفهم كيفية تقاطعها.

يمكن تعريف إدارة بيانات البحث بأنها مجموعة من الممارسات للتعامل مع المعلومات التي تُجمع وتُنشأ في أثناء البحث. فهي «حصيلة ممارسات صغيرة كثيرة تجعل بياناتك أسهل في العثور عليها، وأسهل في فهمها، وأقل عرضة للضياع، وأرجح قابلية للاستخدام في أثناء المشروع أو بعد عشر سنوات». [5] وتشمل هذه الممارسات، دون أن تقتصر عليها، التخطيط لإدارة البيانات، وتوثيقها، وتنظيمها، وتخزينها، ونشرها، وحفظها. [6] والإدارة الفعالة لبيانات البحث عملية مستمرة منظمة ومتسقة مع سياق البحث والممارسات التخصصية.

تدعو مبادئ (FAIR) إلى زيادة قابلية بيانات البحث، والكائنات الرقمية العلمية عمومًا، للعثور عليها والوصول إليها والتشغيل البيئي وإعادة الاستخدام. وتحت مظلة اختصار (FAIR)، صيغت 15 قاعدة لتوجيه تصرفات ناشري البيانات وأمنائها وغيرهم من أصحاب المصلحة. [7] ومن صميم مفهوم (FAIR) انطباقه «على الأنشطة التي يقودها البشر وتلك التي تقودها الآلات على السواء». بهدف تحقيق قابلية التنفيذ الآلي (machine-actionability) إلى أعلى درجة ممكنة أو مناسبة. [8] وفضلاً عن ذلك، فإن (FAIR) ليست ثنائية (أي متوافقة/غير متوافقة)، بل هي طيف

تتاج على امتداده درجات متفاوتة من «التوافق مع (FAIR)» [9]، [10]، [11] ومع أن مبادئ (FAIR) شهدت تنبياً وقبولاً سريعين، فإن للعمل الحالي المرتبط بها اتجاهات كثيرة، منها تطبيقات مختلفة لـ (FAIR) في تقييم الخدمات الداعمة لبيانات (FAIR) وتنفيذها.

البيانات المفتوحة هي ممارسة إتاحة بيانات البحث الأساسية للعموم، بحيث تكون متاحة وقابلة لإعادة الاستخدام بأدنى حد من القيود. وفي إطار التحول الأوسع نحو العلم المفتوح (open science)، [12] أصبحت البيانات المفتوحة على نحو متزايد أمراً تتوقعه الجهات الممولة وواضعو السياسات، وكثيراً ما تُؤطرها القاعدة القائلة «مفتوحة قدر الإمكان، مغلقة قدر الضرورة». ويمكن تعريف البيانات المفتوحة على متصل متدرج، كأن يُستعار مثلاً نظام النجوم الخمس للبيانات المترابطة المفتوحة (Linked Open Data: LOD) الذي وضعه تيم بيرنرز-لي (Tim Berners-Lee)، [13]، [14] فبحسب بيرنرز-لي، فإن الحد الأدنى المطلوب في البيانات المفتوحة أن تحمل ترخيصاً مفتوحاً (مثل (CC0) من المشاع الإبداعي)، لكن لتحقيق قدر أكبر من الانفتاح وإمكانات إعادة الاستخدام، ينبغي أيضاً أن تكون البيانات مقروءة آلياً، وبصيغة غير احتكارية، وأن تستخدم معايير مفتوحة وترتبط ببيانات أخرى لتوفير السياق. وفي هذا النظام، تُجمع النجوم باستيفاء كل معيار خطوة بخطوة. وفي هذه الدرجات العليا من الانفتاح يكون التداخل مع (FAIR) أعمق ما يكون، لأن كليهما يؤكد على الطرق التي يمكن بها جعل المحتوى ذا معنى لدعم إعادة استخدامه من قبل البشر والآلات.

مفاهيم خاطئة شائعة:

ثمة عدد من المفاهيم الخاطئة حول معنى إدارة بيانات البحث ومبادئ (FAIR) والانفتاح. فكثيراً ما يُخلط بين هذه المصطلحات وتُستخدم بالتبادل. ونحاول هنا تفكيك بعض أكثر هذه المفاهيم الخاطئة شيوعاً.

يجب أن تكون بيانات (FAIR) مفتوحة:

كلاً! فمع أن سياسات كثيرة تدعو إلى بيانات متوافقة مع (FAIR) ومفتوحة، فإن الأمرين لا يعينان الشيء نفسه. فيمكن أن تكون البيانات متوافقة مع (FAIR) ومفتوحة معاً، أو أن تكون أحدهما فقط، أو لا هذا ولا ذاك. ومن نقاط قوة مبادئ (FAIR) أنها تسمح بالوصول المضبوط (controlled access)، وهو ما قد يكون مهماً لأنواع معينة من البيانات. وكلاهما أيضاً مقياس متدرج يمكن على امتداده جعل البيانات أو المخرجات الأخرى، مثل الشيفرة البرمجية، متوافقة مع (FAIR) ومفتوحة بدرجات متزايدة [15].

البيانات المفتوحة أكثر فائدة من بيانات (FAIR):

هذان المفهومان ليسا في تنافس؛ فكلهما قيم، وينبغي أن نشجع الباحثين على جعل بياناتهم متوافقة مع (FAIR) ومفتوحة قدر الإمكان. فأكثر البيانات قابلية لإعادة الاستخدام هي البيانات الموثقة جيداً، المتوافقة مع معايير المجتمع العلمي المعني، والخالية من القيود قدر الإمكان لزيادة إمكانات إعادة استخدامها.

كل البيانات المتوافقة مع (FAIR) والمفتوحة عالية الجودة:

لا تعبر بيانات (FAIR) ولا البيانات المفتوحة عن جودة البيانات. فكلهما مجرد مقياس لكيفية إتاحة البيانات. ويمكن لمجموعة بيانات رديئة الجودة أو ملفقة أن تكون متوافقة مع (FAIR) ومفتوحة معًا. ولهذا من المهم إدارة البيانات وتوثيقها جيدًا لتوثيق مصدر البيانات وتاريخها (provenance) والضمانات حول كيفية إنشاء البيانات ومعالجتها، بما يبعث على الثقة. ولكي تكون البيانات ذات أكبر قيمة، ينبغي أن تكون مُدارة جيدًا وأن تُقدّم مع سياق كافٍ يتيح لمن يعيدون استخدامها تقييم ما إذا كانت تفي بأغراضهم.

مبادئ (FAIR) مقصورة على الاتحاد الأوروبي وعلوم الحياة، فلماذا أهتم بها:

مع أن مبادئ (FAIR) انبثقت من ورشة عمل في علوم الحياة في لايدن (Leiden)، فقد صيغت المبادئ عمدًا بمعنى واسع لتطبيق على جميع أنواع البيانات. والواقع أنها تُطبق في سياقات متنوعة؛ فقد وضعت المفوضية الأوروبية مبادئ (FAIR) في صميم مشروعها التجريبي لبيانات البحث إلى جانب البيانات المفتوحة. [16] وخارج أوروبا، لدى الاتحاد الأمريكي للجيوفيزياء (American Geophysical Union: AGU) مشروع بعنوان «تمكين بيانات (FAIR)» (Enabling FAIR Data)، [17] ويدعم مشاع بيانات البحث الأسترالي (Australian Research Data Commons: ARDC) برنامجًا لمبادئ (FAIR). [18]

نمذجة العلاقة بين إدارة بيانات البحث ومبادئ (FAIR) والانفتاح:

كما أوضحنا أعلاه، لكل من إدارة بيانات البحث ومبادئ (FAIR) والانفتاح مواضع تركيز مختلفة. ولا ينبغي أن تُدرج إدارة البيانات ضمن (FAIR) أو الانفتاح، لأنها تتناول ممارسات على امتداد دورة حياة كاملة، ولها فوائد داخلية للباحث والمشروع والمؤسسة لا ترتبط دائمًا بمشاركة البيانات التي تؤكد عليها (FAIR) والانفتاح. وعلى وجه الخصوص، لا تغطي (FAIR) والانفتاح مسائل جودة البيانات، مع أنها حاسمة لإعادة الاستخدام وتدعمها الإدارة والرعاية المناسبان للبيانات على امتداد دورة حياتها. فإدارة بيانات البحث ومبادئ (FAIR) والانفتاح كلها مهمة في حد ذاتها، وينبغي النظر إليها على أنها متكاملة ولكنها متميزة.

يمكن تصور العلاقة بين إدارة بيانات البحث ومبادئ (FAIR) والانفتاح بوصفها ثلاثة أبعاد مستقلة ومتقاطعة، يمتد كل منها على طيف متدرج من الممارسة.

يقرّ النموذج المقترح للعلاقة بين إدارة البيانات ومبادئ (FAIR) والانفتاح بالتفاوت على امتداد الأبعاد الثلاثة جميعها. فالبيانات يمكن أن تكون:

- مُدارة بدرجات متفاوتة، من غير مُدارة إلى مُدارة جيدًا
- مفتوحة بدرجات متفاوتة، من مغلقة تمامًا إلى مفتوحة إلى حد كبير
- متوافقة مع (FAIR) بدرجات متفاوتة، من توافق منخفض إلى توافق مرتفع.

وبوجه عام، تبلغ قيمة البيانات أقصاها عندما يتحقق الانفتاح والتوافق مع (FAIR) كلاهما بدرجة عالية. فالبيانات المصنفة على أنها عالية التوافق مع (FAIR) يفترض أن تكون قد أُدبرت جيدًا، لكنها قد تكون مفتوحة أو مغلقة. وفي حالات أخرى، يمكن أن تجعل البيانات مفتوحة أو متوافقة مع (FAIR) إلى حد ما دون أن تكون مُدارة جيدًا، فينتج عن ذلك بيانات ضعيفة التوثيق وأقل قابلية لإعادة الاستخدام. ولهذا من المهم أن تكون البيانات مُدارة جيدًا أيضًا لدعم المشاركة على نحو ذي معنى وتعزيز إعادة الاستخدام.

الإدارة الجيدة للبيانات شرط سابق ضروري لـ (FAIR) والانفتاح، وهي تتيح إنشاء بيانات صالحة للمشاركة وإعادة الاستخدام. فكثير من القرارات المتخذة في مرحلتي تخطيط البحث وإدارته تؤثر في إمكانية جعل البيانات متوافقة مع (FAIR) و/أو مفتوحة. وقد تشمل هذه القرارات أدوار مشروع البحث ومسؤولياته، واتفاقات الموافقة (consent agreements)، وملكية البيانات واستخدامها المتفق عليهما مع الشركاء، والتراخيص من مالكي البيانات من الأطراف الثالثة، واختيار صيغ البيانات، واختيار مخططات البيانات الوصفية، واصطلاحات التسمية، وإنشاء البيانات الوصفية وتوثيق البيانات أو التقاطعها. وبالانطلاق من أساس من الإدارة الفعالة لبيانات البحث، يستطيع الباحثون وأمناء البيانات بعد ذلك النظر في المستوى المناسب من التوافق مع (FAIR) والانفتاح لكل مجموعة بيانات على حدة، مع مراعاة عوامل مثل نوع المحتوى، وشروط الوصول، وقيود مشروع البحث، والممارسات التخصصية.

درجات إدارة بيانات البحث ومبادئ (FAIR) والانفتاح:

لتوضيح تقاطعات إدارة بيانات البحث ومبادئ (FAIR) والانفتاح وحدودها وقيودها، ناقش أدناه سيناريوهين. ويبيّن هذان السيناريوهان كيف يمكن لكل من هذه الأفكار أن يدعم رعاية أفضل للبيانات في سياقات مختلفة، وما لكل منها من قيود.

بيانات متوافقة جزئيًا مع (FAIR) ومفتوحة لكنها غير قابلة للاستخدام:

من نتائج إدخال سياسات الدوريات اشتراطات لمشاركة البيانات أن عددًا أكبر من مجموعات البيانات صار يُشارك. لكن هذا لا يؤدي دائمًا إلى بيانات قابلة لإعادة الاستخدام. فقد تستوفي مجموعات البيانات المفتوحة معظم متطلبات (FAIR) وتكون مع ذلك غير قابلة للاستخدام عمليًا أو رديئة الجودة. فملف (CSV) منفرد مصحوب بوصف محدود في مستودع بيانات عام يبدو أنه يستوفي كثيرًا من خانات (FAIR) والانفتاح (مثل المعارف الدائمة، والبيانات الوصفية الأساسية، وصيغ الملفات غير الاحتكارية، وغيرها)، لكن محدودية التوثيق تجعل البيانات غير قابلة للاستخدام من دون مزيد من المعلومات عن مصدرها ونشأتها، وشرح المتغيرات، والمنهجية.

وقد تُنشر البيانات أيضًا في صورة رسوم بيانية وجداول بصيغة صور، أو في صورة ملفات تكميلية لا يمكن التعامل معها وإعادة استخدامها مباشرة، مثل ملفات (PDF). ولا يعني هذا أن منشئ البيانات لم يدرها جيدًا، بل يعني أن صيغة قابلة لإعادة الاستخدام لم تُشارك، وكثيرًا ما يعود ذلك إلى سياسة الناشر. ومن الأهمية بمكان أن ننقل مفاهيم (FAIR) والانفتاح وإدارة بيانات البحث بفعالية، حتى يفهم الباحثون القيود المحتملة للملفات التكميلية ويدركوا أن البيانات مخرّج بحثي قيم في حد ذاته. ويجب أن تُشارك البيانات بصيغ قابلة للتحرير ومع توثيق كافٍ يتيح تقييمها وإعادة استخدامها وربما دمجها مع بيانات أخرى.

بيانات نموذجية مغلقة إلى جانب شيفرة متوافقة مع (FAIR) ومفتوحة:

في بعض التخصصات، ومنها الهندسة وعلوم الحاسوب، كثيرًا ما تكون الشيفرة والبرمجيات التي تُطوّر أهم للبحث من البيانات، التي تُستخدم أساسًا لاختبار الشيفرة. وفي هذه التخصصات، من المشكوك فيه إلى أي حد ينبغي أن تُدار البيانات وتُتاح إتاحة مفتوحة. فهذه البيانات، التي كثيرًا ما تسمى بيانات النماذج أو البيانات الاصطناعية (synthetic data)، قد تكون غير مُدارة ومغلقة ولا تلتزم بمبادئ (FAIR)، في حين يمكن أن تكون الشيفرة مُدارة بعناية وموثقة ومتاحة إتاحة مفتوحة. والمرونة في مبادئ (FAIR) تعني أنها قابلة للتطبيق بسهولة على الشيفرة أيضًا، لأن للشيفرة كثيرًا من خصائص البيانات نفسها، بما في ذلك معايير المجتمع العلمي، والمعرفات الدائمة، والترخيص. [19] وهكذا يمكن استخدام مبادئ (FAIR) لإجراء حوار مفيد حول ما يلزم لتحسين شفافية البحث وقابليته لإعادة الإنتاج، سواء اعتمد أساسًا على البيانات أم على الشيفرة.

التوعية بإدارة بيانات البحث ومبادئ (FAIR) والانفتاح والدعوة إليها:

كثيرًا ما لا تكون إدارة بيانات البحث ومشاركتها أولوية عالية عند الحديث مع الباحثين، ومع أن إدارة بيانات البحث ومبادئ (FAIR) والانفتاح تساعد جميعها على تشجيع الممارسات الجيدة، فإن هذا التكاثر في المصطلحات قد يسبب الالتباس أحيانًا. ويلزم تفكير متأن في كيفية استخدام هذه المفاهيم ومتى. والمقترحات المعروضة أدناه خلاصة للقضايا الرئيسية التي أثارها الممارسون في فعاليات متعددة على مدى العامين الماضيين، مستندةً بدرجة كبيرة إلى جلسة نقاش حرة لذوي الاهتمام المشترك (birds of a feather: BoF) عُقدت في مؤتمر «إشراك الباحثين في الإدارة الجيدة للبيانات» (Engaging Researchers in Good Data Management Conference) في كامبريدج عام 2017. [20] وقد تركز النقاش بين أمناء المكتبات وأمناء البيانات والباحثين في هذه الفعالية على كيفية استخدام الممارسين لمبادئ (FAIR) والانفتاح في الدعوة إلى الإدارة الفعالة للبيانات. ونلخص هنا خمس توصيات لاستخدام (FAIR) والانفتاح عند الدعوة إلى إدارة بيانات البحث:

التركيز على (FAIR) عندما يتعذر أن تكون البيانات مفتوحة:

من صعوبات الدعوة إلى البيانات المفتوحة أن الباحثين الذين يعملون مع مشاركين من البشر، ولا سيما من يبحثون في موضوعات حساسة، كثيرًا ما لا يستطيعون مشاركة بيانات بحثهم، وقد لا يرغبون في طلب إذن المشاركين بالمشاركة خشية أن يشي ذلك الناس عن المشاركة في بحثهم. ويمكن أن تكون (FAIR) مفيدة للغاية هنا لأنها لا تشترط سوى إتاحة البيانات الوصفية. ويمكن لهذه البيانات الوصفية أن تشير إلى مجموعات بيانات مغلقة، وإن كان الأمثل أن تكون البيانات متاحة بشكل ما أيضًا. وكون البيانات لا يلزم بالضرورة أن تكون متاحة للعموم يفتح حوارات مع الباحثين لم تكن ممكنة من قبل. وذكر المشاركون في جلسة النقاش أنهم يستخدمون (FAIR) لتشجيع الباحثين على مشاركة جزء من بياناتهم أو تسجيلية من البيانات الوصفية فقط تصف مجموعة البيانات بالتفصيل وتبين خطوات الوصول إليها. وهذا مفيد خاصةً بعد بدء العمل باللائحة العامة لحماية البيانات (General Data Protection Regulation: GDPR)، إذ يمكن القول إن الباحثين صاروا أكثر قلقًا من أي وقت مضى بشأن مشاركة البيانات الشخصية، واستخدام (FAIR) لبدء هذا الحوار يتيح الترويج لإدارة أفضل لبيانات البحث عمومًا.

الاعتراف بالأثر والاستشهاد والمكانة:

عند مناقشة بيانات (FAIR) والبيانات المفتوحة على السواء، أثار كثير من المشاركين قضايا تتعلق بالاستشهادات والأثر والمكانة وتقييم الباحثين. ومع أن هذه القضايا قد لا تكون محور تركيز إدارة بيانات البحث، فمن المهم الاعتراف بالضغوط الواقعة على الباحثين لنشر أبحاثهم في دورية مرموقة، والوضوح بشأن الكيفية التي يمكن بها لمشاركة البيانات على نحو متوافق مع (FAIR) ومفتوح أن تساعد على دعم ذلك. وهنا يمكن لعنصر «قابلية العثور» في (FAIR) أن يساعد على تشجيع الإدارة الجيدة لبيانات البحث بمخاطبة رغبة الباحثين في جعل البيانات أكثر ظهورًا، مما يحفزهم على تضمين بيانات وصفية كافية في مجموعة البيانات لجعلها قابلة للعثور عليها. ويمكن أيضًا ربط التأكيد على قابلية إعادة الاستخدام في (FAIR) بأجندة الأثر، بتشجيع استخدام مجموعات البيانات المشاركة استخدامًا فعليًا في سياقات أخرى بدلًا من مجرد الإشارة إليها.

استخدام المصطلحات للمساعدة لا للتعمية أو التخويف:

معظم الباحثين على دراية من قبل بمفهوم «الانفتاح» من خلال الوصول المفتوح (open access: OA)، وهذا يمكن أن يساعد على بدء الحوار مع الباحثين الذين لم يفكروا في مشاركة البيانات من قبل. لكن يلزم موازنة ذلك مع ارتباط الوصول المفتوح بالامتثال، وفي المملكة المتحدة بإطار التميز البحثي (Research Excellence Framework: REF). فالبيانات المفتوحة مفهوم بسيط نسبيًا، على الأقل في ظاهره، ويمكن فهمه بسرعة، وإن كان تطبيقه صعبًا في بعض الحالات. وفي تطبيق مشاركة البيانات تحديدًا يمكن أن تكون (FAIR) مفيدة بوصفها أداة للتوعية والدعوة. فمع أن (FAIR) تضيف اختصارًا آخر يحتاج إلى شرح، فإنه يمكن بعد ذلك استخدامها بفعالية لإرشاد الباحثين خطوة خطوة عبر ما يلزم لجعل بياناتهم متاحة ومفهومة. وتتضمن مبادئ (FAIR) الخمسة عشر عدة نقاط عمل واضحة، مثل الحصول على معرف دائم، وإسناد ترخيص للاستخدام، وتوفير البيانات الوصفية على الإنترنت.

مراعاة المستخدمين النهائيين للبيانات:

من الموضوعات المتكررة السؤال عما إذا كان الأهم تعظيم كمية البيانات المتاحة أم امتلاك عدد أقل من مجموعات البيانات عالية الجودة. فعندما يكون الانفتاح هو محور التركيز الوحيد، ثمة خطر أن تكون كثير من مجموعات البيانات المشاركة منخفضة الجودة، ببيانات وصفية ضعيفة، بحيث لا يستطيع فهم البيانات إلا الأقران المباشرون للباحث. وعلى العكس من ذلك، قد تؤدي الدعوة إلى (FAIR) إلى مشاركة عدد أقل من مجموعات البيانات بسبب المتطلبات المتزايدة النابعة من المبادئ، لكنها قد تكون أغنى وأيسر فهمًا خارج التخصص. ولذا ينبغي لأي حوار حول مشاركة البيانات أن يركز على الجمهور الذي يسعى الباحث إلى الوصول إليه، وعلى الاستخدامات الممكنة لبياناته، إلى جانب جعلها متوافقة مع (FAIR) ومفتوحة. ومن المهم الاعتراف بالجهد الذي تنطوي عليه إدارة البيانات ومشاركتها، حتى يمكن إصدار أحكام معقولة حول متى وأين يُطبق ذلك، كما يبين مثال بيانات النماذج أعلاه.

إبقاء البحث والباحث في صميم الرسالة:

سواء استخدمنا مصطلحات إدارة بيانات البحث أو (FAIR) أو الانفتاح، فمن السهل جدًا على من يدعون منا إلى إدارة البيانات ومشاركتها أن ينشغلوا باهتمامهم بالبيانات وينسوا ما يهم الباحث: بحثه! فنحن لا ندعو إلى إدارة البيانات ومشاركتها لذاتهم، ومع أن الغايات النهائية تتنوع من قابلية إعادة الإنتاج، إلى إعادة الاستخدام عبر التخصصات،

إلى التطبيق من قبل الممارسين، فإن الاستخدام الذي يمكن أن توظّف فيه تلك البيانات ينبغي أن يكون في صميم أنشطتنا. ولذا، فمع أن الاختصارات والمصطلحات المألوفة قد تكون وسيلة مختصرة مفيدة لوصف ممارسات إدارة البيانات المعقدة أحياناً التي نشجعها، فلا ينبغي أبداً أن تحل محل التركيز على تحسين كفاءة عملية البحث وشموليّتها والمساعدة على الإجابة عن أسئلة بحثية جديدة.

الاستنتاجات:

إدارة بيانات البحث ومبادئ (FAIR) والانفتاح كلها مهمة في حد ذاتها، وينبغي النظر إليها على أنها مفاهيم متكاملة ولكنها متميزة. فالثلاثة جميعاً توجد على طيف ويتقاطع بعضها مع بعض: إذ يمكن أن تُدار البيانات بدرجات متفاوتة، وأن تكون متوافقة مع (FAIR) ومفتوحة بقدر أكبر أو أقل. وينبغي أن ننظر إلى كل منها على أنه مستوى من النضج يُشجّع فيه الباحثون على جعل بياناتهم أفضل إدارة وأكثر توافقاً مع (FAIR) وانفتاحاً، حتى تكون في نهاية المطاف أكثر فائدة. فإدارة البيانات هي الشرط السابق الضروري لتمكين بيانات (FAIR) والبيانات المفتوحة، وفي المقابل يمكن لهذه المبادئ أن تساعد على الدعوة إلى ممارسات جيدة في إدارة البيانات.

ولا يكفي بالضرورة أن تكون البيانات متوافقة مع (FAIR) ومفتوحة. فقد صُمم الإنترنت آليةً لمشاركة المحتوى بين مواقع موثوقة ذات سلطة. وصار بإمكان أي شخص أن يكون منشئاً للبيانات وناشراً لها على الإنترنت. وثمة ضوابط قليلة تساعد المستخدمين على معرفة البيانات التي يمكن الوثوق بها، ومن هنا تأتي أهمية الرعاية المهنية (professional curation) في المستودعات المعتمدة لضمان رعاية البيانات رعاية فعالة وبقائها متاحة على المدى الطويل.

تشمل الحدود والتقاطعات بين إدارة بيانات البحث ومبادئ (FAIR) والانفتاح عناصر مهمة معرضة لخطر الإغفال إذا ركزنا على مفهوم واحد فقط. فبيانات (FAIR) التي تحظى برعاية سليمة لها إمكانات كبيرة لإعادة الاستخدام، لكن إذا أمكن أيضاً إتاحتها بيانات مفتوحة، فإن إمكانات إعادة الاستخدام هذه تزداد. وبالمثل، إذا عُرّفت البيانات المفتوحة تعريفاً فريداً بحيث يمكن اكتشافها، وخضعت لرعاية مهنية على المدى الطويل، فستزداد احتمالية إعادة استخدامها وعمق هذا الاستخدام. وينبغي أن ندعو إلى أن تكون البيانات متوافقة مع (FAIR) ومفتوحة قدر الإمكان، مستخدمين هذه المبادئ للمساعدة على غرس ممارسات جيدة في إدارة البيانات منذ البداية. فالكل أعظم من مجموع أجزائه.

المراجع:

1. "5-star Open Data," James G. Kim and Michael Hausenblas, <https://5stardata.info/en/> (accessed April 5, 2019).
2. "FAIR Data Maturity Model WG," Research Data Alliance, <https://www.rd-alliance.org/groups/fair-data-maturity-model-wg> (accessed April 5, 2019).
3. "Introduction: Open Science Training Handbook," FOSTER, <https://book.fosteropenscience.eu/en/01Introduction/> (accessed April 5, 2019).
4. "Design Issues: Linked Data," Tim Berners Lee, <https://www.w3.org/DesignIssues/LinkedData.html> (accessed April 5, 2019).

5. "Open Research Data the New Norm in H2020," OpenAIRE, <https://www.openaire.eu/open-research-data-the-new-norm-in-h2020> (accessed April 5, 2019).
6. "Enabling FAIR data project," Coalition for Publishing Data in the Earth and Space Sciences, <http://www.copdess.org/enabling-fair-data-project/> (accessed April 5, 2019).
7. "The FAIR data principles," Australian National Data Service, <https://www.andis.org.au/working-with-data/fairdata> (accessed April 5, 2019).
8. "Engaging Researchers in Good Data Management Conference 2017," University of Cambridge, <https://www.repository.cam.ac.uk/handle/1810/270234> (accessed April 5, 2019).
9. FORCE11, "The FAIR Data Principles," <https://www.force11.org/group/fairgroup/fairprinciples> (accessed April 5, 2019).
10. OECD, Principles and Guidelines for Access to Research Data from Public Funding, 12 April 2007; DOI: <https://doi.org/10.1787/9789264034020-en-fr> (accessed April 5, 2019).
11. Foreign and Commonwealth Office, G8 Science Ministers Statement, 13 June 2013, <https://www.gov.uk/government/news/g8-science-ministers-statement> (accessed April 5, 2019).
12. Sarah Jones, "Open data, FAIR data and RDM: the ugly duckling," March 2018; DOI: <https://doi.org/10.1145/3158344> (accessed April 5, 2019).
13. Kristin Briney, Data Management for Researchers (Exeter, UK: Pelagic Publishing, 2015).
14. Louise Corti, Veerle Van den Eynden, Libby Bishop and Matthew Woollard, Managing and Sharing Research Data (London: SAGE, 2014).
15. FORCE11, "The FAIR Data Principles".
16. Mark D. Wilkinson et al., "The FAIR Guiding Principles for scientific data management and stewardship," Scientific Data 3 (2016): 3; DOI: <https://doi.org/10.1038/sdata.2016.18> (accessed April 5, 2019).
17. Barend Mons et al., "Cloudy, increasingly FAIR; Revisiting the FAIR Data guiding principles for the European Open Science Cloud," Information Services & Use 37, no. 1 (2017): 49-56; DOI: <https://doi.org/10.3233/ISU-170824> (accessed April 5, 2019).
18. Mark D. Wilkinson et al., "A design framework and exemplar metrics for FAIRness," Scientific Data 5 (2018); DOI: <https://doi.org/10.1038/sdata.2018.118> (accessed April 5, 2019).

19. European Commission Expert Group on FAIR Data, Turning FAIR into reality, 26 November 2018, <https://publications.europa.eu/en/publication-detail/-/publication/7769a148-f1f6-11e8-9982-01aa75ed71a1> (accessed April 5, 2019).
20. Neil Chue Hong and Daniel S. Katz, "FAIR Enough? Can we (already) benefit from applying the FAIR data principles to software? (Version 2)," December 2018, <https://doi.org/10.6084/m9.figshare.7449239.v2> (accessed April 5, 2019).