عشر قواعد بسيطة تساعد المؤسسات على دعم مشاركة بيانات البحث

روبن شامبيو، أنتوني سولومونيدس، ماريسا كونتي، سفيتلانا روجيفسكي، جيمي فوونغ، ديفيد إيه. دور، إليزابيث زامبينو، آدم ويلكوكس، ماثيو بي. كارسون، كريستي هولمز



الملخص:

يقدّم هذا المقال إرشادات للمؤسسات البحثية لبناء قدرتها على دعم مشاركة بيانات البحث، انطلاقًا من أن نجاح الباحث في المشاركة يتوقف على بيئته المؤسسية. وتعرض القواعد الثلاث الأولى أدوات تأسيسية: استخدام نموذج للنضج لتقييم القدرات الحالية، والاستفادة من الشخصيات الافتراضية وقصص المستخدمين لفهم الدوافع والتحديات المحلية، وتعزيز التواصل والتعاون. وتتناول القواعد من الرابعة إلى العاشرة سبعة مجالات مستمدة من نموذج «مشاركة بيانات البحث: نموذج نضج للقدرة المؤسسية»، هي: الحوكمة، والإجراءات وسير العمل، والقيم المؤسسية، والبنية التحتية، وتدريب القوى العاملة، وجودة البيانات وإعادة استخدامها، وأخلاقيات البيانات. وتُرفق بكل قاعدة أسئلة نموذجية قد يطرحها العاملون في البحث، مع أمثلة لموارد ومبادرات قائمة. ويخلص المؤلفون إلى أن هذه المجالات مترابطة، وأن مستويات النضج طموحات توجّه التخطيط، ويدعون مجتمعات التخصصات الأخرى إلى مراجعة النموذج وتكييفه.

الكلمات المفتاحية: مشاركة بيانات البحث، نماذج النضج، حوكمة البيانات، البنية التحتية للبيانات، أخلاقيات البيانات

مقدمة:

يعتمد الاكتشاف العلمي على الوصول إلى البيانات وإلى المعرفة التي تتيحها هذه البيانات. ويتزايد الاعتراف بمشاركة بيانات البحث (research data sharing) أولويةً ينبغي للمؤسسات دعمها لضمان نجاح إجراء البحوث. وتنص المعاهد الوطنية للصحة (National Institutes of Health) على أن «مشاركة البيانات تمكّن الباحثين من اختبار صحة نتائج البحث اختبارًا صارمًا، وتعزيز التحليلات من خلال مجموعات البيانات المدمجة، وإعادة استخدام البيانات التي يصعب توليدها، واستكشاف آفاق جديدة للاكتشاف» [1]. وفي المقابل، تزداد في غياب مشاركة

البيانات المخاطرُ المتعلقة بمتانة النتائج وصرامتها وقابليتها للتكرار، وتتضاءل إمكانات البيانات القيّمة. ولهذه الأسباب وغيرها، تُعد قدرة المؤسسة على مشاركة البيانات موضوعًا بالغ الأهمية ينبغي للمؤسسات فحصه ومناقشته والنهوض به.

وكثيرًا ما تُناقش الدعوة إلى مشاركة البيانات ودعمها مع التركيز على فهم ممارسات الباحثين الأفراد أو المجتمعات العلمية ودعمها [2,3]. غير أن قدرة الباحث على المشاركة بنجاح في مشاركة سليمة للبيانات والإفادة منها تعتمد على بيئته المؤسسية، وتحديدًا على قدرة المؤسسة على مشاركة البيانات. فعلى سبيل المثال، تكون مشاركة البيانات أيسر وأكثر إنصافًا حين تكون العمليات والإجراءات المؤسسية راسخة وموثّقة، وحين يتمكن العاملون في البحث من الوصول إلى موارد مركزية للتدريب والبنية التحتية.

ويمتد أثر الطريقة التي تتعامل بها المؤسسة مع مشاركة البيانات وتدعمها إلى ما هو أبعد من نجاح باحثيها. فالمؤسسات التي تشارك البيانات يمكنها المشاركة في مبادرات مبتكرة واسعة النطاق والسعي إلى فرص تمويل جديدة. والجامعات التي تسهم في ابتكارات مشاركة البيانات، بما في ذلك أفضل الممارسات والإرشادات الأخلاقية، يمكنها أن تحقق آثارًا أعمق للمجتمعات التي تدعمها وتخدمها. والمؤسسات التي تستثمر عن قصد في مشاركة بيانات البحث تكون أكثر استعدادًا للامتثال لمشهد السياسات المتطور. ومن ثَمّ فإن قدرة المؤسسة على مشاركة البيانات عاملٌ في النجاح العلمي ومؤشرٌ عليه في آن واحد.

وتيسّر هذه الورقة نقاشًا عامًا للمؤسسات حول القدرة على مشاركة البيانات. فنحن نعرّف بفائدة نماذج النضج (القاعدة 1)، ونناقش أهمية فهم المستخدمين وحالات الاستخدام (القاعدة 2)، ونؤكد قيمة التواصل والتعاون (القاعدة 3). وللحث على الاعتراف بطائفة من الجوانب الرئيسة وفحصها، نعرض أيضًا قواعد تبرز 7 عناصر مؤسسية تمكّن مشاركة البيانات أو تعيقها (القواعد من 4 إلى 10)، استنادًا إلى نموذج نضج حديث [4]. وإجمالًا، نهدف إلى لفت الانتباه إلى مختلف الأنظمة والتقنيات والسياسات والممارسات والقيم والأشخاص الذين يؤثرون في فاعلية المؤسسة في مشاركة البيانات. ويمكن لهذه القواعد أن توجّه المؤسسة عبر الأنشطة الاستراتيجية والعملية الرئيسة المتعلقة ببيانات البحث، من تكوين فريق وتقييم الاحتياجات عبر الجوانب الرئيسة، إلى تطوير الهياكل الرسمية والموارد والخدمات اللازمة أو تعزيزها لتتمكن المؤسسة من تيسير مشاركة البيانات بنجاح. ومع أن هذه الورقة كُتبت للمؤسسات، فإنها تهدف إلى أن تعكس تجارب مجتمعات البحث المحلية واحتياجاتها. ولذلك أدرجنا في عدد من القواعد أسئلة رئيسة قد يطرحها العاملون في البحث، وينبغي للمؤسسات أن تكون مستعدة للإجابة عنها بمعلومات يمكن العمل بها.

القاعدة 1: استخدم نموذجًا للنضج لفهم الخصائص المؤسسية المؤثرة في مشاركة البيانات:

تحتاج المؤسسات إلى فهم قدراتها الأساسية على مشاركة البيانات، ويمكن لنموذج النضج (maturity model) أن ييسّر هذا التقييم. ونموذج النضج إطار لوصف العمليات والهياكل والتقنيات والثقافة والأشخاص الذين يدعمون تنامي الفاعلية في مجال محدد، مثل إدارة البيانات أو تقنية المعلومات البحثية، ولتقييمهم [5,6]. وتُستخدم نماذج النضج لتحديد نقاط القوة والضعف وإعداد خطط التحسين، إذ تصف العوامل أو المجالات ذات الصلة على مقياس متدرج للقدرة. وبذلك تستطيع المؤسسة تحديد موقعها في النموذج وتقييم قدراتها الحالية في كل مجال. وتفيد هذه الخاصية بوجه خاص المؤسسات التي بدأت حديثًا التخطيط لمشاركة البيانات على المستوى المؤسسي.

ويصف نموذج «مشاركة بيانات البحث: نموذج نضج للقدرة المؤسسية» (Research Data Sharing: A Maturity Model for Organizational Capacity) سبعة مجالات مترابطة: الحوكمة، والعمليات والإجراءات، والثقافة المؤسسية، والبنية التحتية، وتطوير القوى العاملة، وجودة البيانات وإعادة استخدامها، وممارسات أخلاقيات البيانات [4]. وتتدرج مستويات كل مجال من المستوى 1، الذي يتسم بنقص الموارد أو العمل المركّز، إلى المستوى 5، الذي يتسم بقدرات مستقرة ومتحسنة باستمرار. ومع أن نموذج النضج هذا شامل، فإنه لا يُقصد به أن يكون إلزاميًا. وهدفنا أن نستوعب تنوع المؤسسات المهتمة بمشاركة البيانات أو باستخدام البيانات المشارَكة، مع لفت الانتباه إلى العمليات والهياكل والتقنية والعناصر الثقافية والأشخاص الذين لا غنى عنهم لتحقيق ذلك. ويمكن للمؤسسات استكشاف كل مجال من المستوى 1 إلى المستوى 5 لفهم ما تبدو عليه القدرة المتزايدة، ولتحديد درجة نضجها الحالية، التي يُرجَّح أن تختلف من مجال إلى آخر. ويمكن أن يكون «مشاركة بيانات البحث: نموذج نضج للقدرة المؤسسية» أيضًا أداةً لتوجيه تحليل الفجوات، بما يساعد المؤسسات على تحديد استثمارات وتدخلات محددة لرفع أدائها في مشاركة البيانات.

القاعدة 2: استفد من الشخصيات الافتراضية وقصص المستخدمين لفهم الدوافع والتحديات المحلية فهمًا أفضل:

يتطلب النهوض بمشاركة البيانات على المستوى المؤسسي تقييمًا شاملًا لأصحاب المصلحة. والشخصيات الافتراضية للمستخدمين (user personas) أداة مفيدة لتحديد طائفة الأدوار والخبرات المنخرطة في مشاركة البيانات وفهمها. والشخصية الافتراضية ملف تعريفي مركّب للمستخدم يصف سمات نمطية، مثل الدوافع والصعوبات والتطوير المهني والمواقف التقنية. فعلى سبيل المثال، تتضمن «الشخصيات الافتراضية للقوى العاملة الانتقالية» (Personas for the Translational Workforce) طائفة من الأدوار ذات الصلة بمشاركة البيانات، مثل محلل البيانات، والإحصائي الحيوي، وأمين المكتبة، ومدير مجلس المراجعة المؤسسية (Institutional Review Board, IRB)، ومدير البحث [7]. ويمكن للتأمل المنهجي في أدوار العاملين، بمساعدة هذه الشخصيات الافتراضية، أن يساعد المؤسسات على تقييم التدريب والموارد وأشكال الدعم الأخرى لمشاركة البيانات وتنفيذها.

وقصص المستخدمين (user stories) أداة فعالة أخرى لبناء معرفة قابلة للتنفيذ عن الأشخاص والأنشطة المنخرطين في مشاركة البيانات، والاستجابة لاحتياجاتهم. وتتبع قصص المستخدمين عادةً صيغة معيارية، مثل قالب (Connextra) [8]، الذي يقدّم إطارًا بسيطًا لكتابة نص موجز يعبّر عن قصة المستخدم: بصفتي ___ [مَن/الشخصية]، أريد ___ [ماذا/المتطلب]، كي ___ [لماذا/النتيجة]. ويساعد إدراج قصص المستخدمين على توفير سياق إضافي عن المستخدمين النهائيين ووجهات نظرهم. ويمكن إنشاء الشخصيات الافتراضية وقصص المستخدمين أو تحديثها حسب الحاجة لتأطير الاستثمارات اللازمة لبناء القدرة المؤسسية على مشاركة البيانات وترتيب أولوياتها على نحو وثيق الصلة بالسياق المحلي وشامل (الجدول 1).

الجدول 1. قصص المستخدمين في مشاركة البيانات:

قصص المستخدمين في مشاركة البيانات
■ بصفتي باحثًا (RESEARCHER) لديه بيانات ناتجة عن مشروع ممول بمنحة، أريد الامتثال لمتطلبات مشاركة البيانات لدى الجهة الممولة لي كي أتمكن من مواصلة بحثي.

■ بصفتي قيّمًا على البيانات (DATA CURATOR)، أحتاج إلى فرص للتطوير المهني ومواد مرجعية لتقييم مجموعات البيانات والعناية بها، لأضمن قدرتي على استخدام أفضل ممارسات البيانات.
■ بصفتي مدير بحث (RESEARCH ADMINISTRATOR)، أحتاج إلى التأكد من أن خطط إدارة بيانات البحث ومشاركتها تتوافق مع متطلبات الجهات الممولة.
■ بصفتي مختصًا في تقنية المعلومات (IT PROFESSIONAL)، أحتاج إلى فهم متطلبات الأنظمة الفدرالية، وكيف يمكن لبنيتنا التحتية أن تلبي على أفضل وجه احتياجات الباحثين الذين يولّدون بيانات حساسة أو خاصة.
■ بصفتي أمين مكتبة (LIBRARIAN)، أريد التعرف على سياسات مشاركة البيانات ومسارات عملها لدعم الباحثين.
■ بصفتي جهة دعم للحوسبة البحثية (RESEARCH COMPUTING SUPPORT)، أحتاج إلى البقاء على اطلاع على أحدث المتطلبات والموارد التي يمكن أن تدعم مسارات عمل مشاركة بيانات البحث في حرمنا الجامعي.
■ بصفتي قائدًا مؤسسيًا (ORGANIZATIONAL LEADER)، أريد معرفة الموارد والخدمات اللازمة كي نتمكن من وضع الميزانية والتخطيط لها.
■ بصفتي مدير مستودع (REPOSITORY MANAGER)، أريد فهم الأولويات المتعلقة بإيداع البيانات، ولا سيما أين ينبغي إيداع أنواع معينة من البيانات، وكيف تُستوفى أفضل الممارسات لجعل البحث قابلًا للاكتشاف، مع صون اعتبارات الخصوصية أو الأمن.
■ بصفتي شريكًا مجتمعيًا (COMMUNITY-BASED PARTNER) لمركز السرطان، أريد الوصول إلى الاكتشافات البحثية للمركز كي أتمكن من متابعة أحدث النتائج في المجالات التي تؤثر في أسرتي ومجتمعي.
■ بصفتي فردًا من الجمهور (MEMBER OF THE PUBLIC)، أريد الوصول إلى نتائج البحوث التي تموّلها الضرائب التي أدفعها، كي تزداد المساءلة والشفافية في البحث.

القاعدة 3: عزّز التواصل والتعاون للنهوض بمشاركة البيانات:

التواصل والتعاون الفعالان أساسيان للبحث الحديث. ونظرًا إلى النطاق المؤسسي الشامل للأنظمة والتقنيات والسياسات والممارسات والقيم والأشخاص المعنيين، فإن الاستثمار في التواصل والتعاون يرعى قدرة المؤسسة على مشاركة البيانات. ويوفر مجال علم العمل الجماعي (team science) أساسًا عمليًا لهذا العمل، ويتضمن طائفة واسعة من الموارد. فعلى سبيل المثال، جمع مركز البحث والتميز والتنوع في علم العمل الجماعي (Center for Research, Excellence and Diversity in Team Science, CREDITS) مجموعة من موارد علم العمل الجماعي في «مجموعة أدوات التعاون الشامل» (Inclusive Collaboration Toolkit) [9]. وتقدّم المجموعة معارف وأدوات لتعظيم كفاءة أنشطة علم العمل الجماعي الشاملة وفاعليتها. و(COALESCE) أداة مجانية عبر الإنترنت للتدريب على علم العمل الجماعي، تقدّم مقاطع فيديو قصيرة وأنشطة تغطي موضوعات مثل اتخاذ القرار التعاوني مع المجتمعات [10]. ويمكن لهذه الموارد وغيرها [11] من مجال علم العمل الجماعي أن تساعد المؤسسات على فهم السياقات التعاونية التي تُولَّد فيها البيانات وتُشارَك، وعلى تعزيز التواصل والتعاون الفعالين، وبناء الثقة بين الشركاء الكثيرين المنخرطين في تطوير الموارد والخدمات للمؤسسة وتقديمها.

وتصف القواعد من 1 إلى 3 مناهج وأدوات تأسيسية لمساعدة المؤسسات على فهم قدرتها على مشاركة البيانات وتنميتها، ولا سيما فيما يتعلق بالعوامل التي تناقشها القواعد من 4 إلى 10. وتُرفق بكل قاعدة أسئلة نموذجية يُرجَّح أن يطرحها العاملون في البحث على مؤسستهم (الأطر 1-7).

القاعدة 4: طوّر هياكل الحوكمة ومكّن القادة وأصحاب المصلحة من قيادة السياسات:

ترسي الحوكمة (governance) عمل المؤسسة في مشاركة البيانات من خلال وضع السياسات والتعريف بها، تحدد التوقعات، وتقدّم إرشادات متسقة للاعتبارات الروتينية، وتعكس احتياجات المؤسسة وقدراتها الفريدة. ومن الناحية المثالية، ينبغي أن يتضمن هيكل الحوكمة مشاركة فعلية من قادة خاضعين للمساءلة ومن أطراف داخلية وخارجية عبر دورة حياة البيانات [6]. وقد يشمل ذلك كبار مسؤولي البحث، وباحثين من تخصصات متنوعة، وأمناء مكتبات، ومختصين في نقل التقنية، ومديري الشؤون التنظيمية، والمشاركين في البحث وشركاءه، وقادة ذوي خبرة تقنية ذات صلة. وتقلل الحوكمة الشاملة والفعالة من اتخاذ القرارات المرتجلة، والجهود المتكررة، والأخطاء التي كثيرًا ما تنجم عنها، مما يجعل مشاركة البيانات أيسر وأكثر كفاءة للباحثين. كما تقلل الحوكمة الشاملة والفعالة من المخاطر المؤسسية، كتلك المتعلقة بالمشاركة الملائمة لأنواع معينة من البيانات (مثل البيانات الحساسة أو البيانات التي يُتصور أن لها إمكانات تجارية)، وتصبح المؤسسة في وضع أفضل للاستجابة للفرص والمتطلبات الجديدة بسرعة ومرونة.

ومن المناهج الفعالة لإرساء هيكل الحوكمة تطوير مصفوفة (RACI)، تُربط فيها الأدوار ومستويات صلاحياتها بوظائف محددة. وتوصف المستويات الأربعة بأنها: المسؤول عن التنفيذ (Responsible) (مَن يؤدي العمل)، والمساءَل (Accountable) (مَن يضمن إنجاز العمل وفق المحدد)، والمستشار (Consulted) (مَن يؤثر في طريقة أداء العمل)، والمُبلَّغ (Informed) (مَن يحتاج إلى المعرفة، حتى لو لم يكن له تأثير مباشر في مضمون العمل). وتوفر (RACI) وضوحًا يضمن تغطية جميع الوظائف، وعدم إثقال الأدوار أو تعارضها، ووجود مسار من الاستشارة والمسؤولية إلى المساءلة.

وتتضمن هياكل الحوكمة المؤسسية على نحو متزايد وحدات شاملة وتعاونية لخدمات بيانات البحث، تتولى الإشراف وتشمل إسهامات من الوحدات الرئيسة والمسؤولين عن إدارة البيانات في الحرم الجامعي، مثل مكاتب البحث، وتقنية المعلومات، والحوسبة البحثية، والمكتبات. ومن الأمثلة البارزة مجموعة خدمات إدارة بيانات البحث في جامعة كورنيل (Cornell University's Research Data Management Service Group, RDMSG) وإدارة بيانات البحث في هارفارد (Research Data Management @Harvard) [12,13].

الإطار 1. أسئلة العاملين في البحث عن الحوكمة:

- ما الوحدات التي تشرف على إجراءات مشاركة البيانات، وكيف تُتخذ القرارات؟
- هل توجد وثائق يمكنني استخدامها لإطلاع مختبري على متطلبات المؤسسة لمشاركة البيانات؟
- إلى أين يمكنني التوجه لطرح أسئلة عن مشاركة البيانات؟
- ما هياكل الحوكمة القائمة للإشراف على البيانات الخاصة أو الحساسة؟

القاعدة 5: وفّر مسارات عمل واضحة لمشاركة البيانات وإجراءات خطوة بخطوة:

تعتمد مشاركة البيانات الفعالة والأخلاقية والمستدامة على مسارات عمل وإجراءات شفافة تعالج الاعتبارات التقنية والاجتماعية واعتبارات الحوكمة. ويتطلب بناء هذه العمليات وصيانتها فهمًا مستنيرًا تمامًا لمشهد مشاركة البيانات في المؤسسة واستراتيجيتها العامة، بما في ذلك الأدوار والاحتياجات الرئيسة لمالكي البيانات والقائمين

على رعايتها ومقدّمي الخدمات والجهات المعيدة لاستخدامها [14]. وستتطلب أنواع البيانات المختلفة مناهج محددة، وينبغي لمسارات العمل والإجراءات الموثقة توثيقًا جيدًا أن توفر مسارات واضحة للتعامل مع المتطلبات التشغيلية والتقنية والتنظيمية.

وتتطلب مشاركة البيانات الناجحة تحديد الأنشطة ومواءمتها، ووضع خطط عمل واضحة لدعم هذه الأنشطة. وتراعي عملية التخطيط هذه الاعتبارات الزمنية والاعتماديات، وكثيرًا ما تكشف عن عمليات موافقة تكرارية (مثل موافقات مجلس المراجعة المؤسسية واتفاقيات استخدام البيانات). وينبغي للمؤسسات أن تقدّم إرشادات واضحة توجّه الامتثال لسياسات بعينها، وأن تنظر في إعداد دليل لإجراءات التشغيل المعيارية (standard operating procedures, SOP) لدعم الموظفين والقادة المسؤولين عن الدعم المركزي الشفاف والروتيني. ويمكن أن يتضمن هذا الدليل عمليات مثل التواصل والرسائل المعتمدة، وأدوار الوحدات المتعاونة ومسؤولياتها، ومسارات العمل المثلى للأنشطة الشائعة، وعمليات البيانات الحساسة أو الحفظ طويل الأمد، وغير ذلك. وتتيح الوثائق الناتجة لفرق البحث ومقدّمي الخدمات المؤسسية إدارة الخطوات الحرجة والحساسة زمنيًا، كتلك المتعلقة بإزالة المعرّفات الشخصية (de-identification) أو نقل كميات كبيرة من البيانات، على نحو شفاف ويُعتمد عليه.

الإطار 2. أسئلة العاملين في البحث عن الإجراءات:

- أين أسجّل لتلقي التحديثات عن مشاركة البيانات؟
- أين يمكنني معرفة المزيد عن أفضل ممارسات مشاركة البيانات التي ينبغي لي اتباعها؟
- كيف أضمن أنني مستعد للامتثال لسياسات الجهات الممولة لي؟
- أحتاج إلى من يراجع خطة إدارة البيانات ومشاركتها قبل أن أقدّم طلب المنحة؛ فمن يمكنه المساعدة؟

القاعدة 6: عزّز القيم المؤسسية لمشاركة البيانات ورسّخها:

تعتمد المشاركة المؤسسية الشاملة في مشاركة مستدامة ومثمرة للبيانات على قيم المؤسسة ويُعبَّر عنها من خلالها. ونرى أن الثقافة المؤسسية لدى المؤسسة فيما يخص مشاركة البيانات تشمل كيف يفسّر القادة والباحثون مشاركة البيانات عمومًا، وكيف تعبّر أنظمة المكافآت فيها عن هذه المواقف، وكيف تتعامل مع قرارات المشاركة في فرص مشاركة البيانات الجديدة وأفضل ممارساتها.

وفي الوقت الراهن، قليلة هي المؤسسات البحثية التي تجسّد تمامًا الرؤية التي يطرحها أعلى مستويات النضج في هذا المجال، حيث: تقدّر المؤسسة مشاركة البيانات تقديرًا عاليًا وعلنيًا؛ ويُعترف بالباحثين على مشاركتهم من خلال عمليات مثل الترقية والتثبيت الوظيفي؛ وتُدرس باستمرار طرق تحديد إسهامات مشاركة البيانات وقياسها؛ وتُعد الموارد اللازمة للانخراط في فرص مشاركة البيانات الجديدة وأفضل ممارساتها استثمارات ضرورية. ومع ذلك، نستلهم أعمالًا سابقة تصف أهمية الثقافة في تيسير مشاركة البيانات والعلم المفتوح [15,16].

ولا تحتاج المؤسسات التي تشرع في هذه الرحلة إلى البدء من الصفر. فقد وضعت أطراف كثيرة، منها الجهات الممولة والهيئات التنظيمية والمكتبات والمجتمعات العلمية، إرشادات وتصورت نماذج مبتكرة لبناء ثقافة لمشاركة البيانات. فعلى سبيل المثال، يقترح بيوّوار (Piwowar) وزملاؤه أن يشجع رؤساء الأقسام أعضاء هيئة التدريس صراحةً على رصد كيفية إعادة استخدام بياناتهم ولماذا، كي يمكن وصف ذلك ومكافأته في قرارات التعيين والترقية [17]. ويوصون أيضًا بأن تتضمن المناهج التمهيدية للبحث مخرجات تعلم وتعليمًا يتعلقان بمشاركة البيانات. ويدعو وود-

تشارلسون (Wood-Charlson) وزملاؤه إلى الاحتفاء بالبيانات المتوافقة مع مبادئ (FAIR) (Findable Accessible Interoperable Reusable)، أي قابلة للعثور عليها، وقابلة للوصول، وقابلة للتشغيل البيني، وقابلة لإعادة الاستخدام) والبيانات القابلة لإعادة الإنتاج، ويقدّمون عدة مقترحات قابلة للتوسع، مثل استخدام قوائم تحقق (FAIR) حين يغادر الطلاب وباحثو ما بعد الدكتوراه مختبرًا أو برنامجًا بحثيًا [18]. وعلاوة على ذلك، ستطوّر مدونة الاستشهاد بالبيانات العالمية المفتوحة (Open Global Data Citation Corpus)، المعلَن عنها حديثًا، مجمّعًا موثوقًا ومتاحًا على نحو مفتوح لجميع الإحالات إلى بيانات البحث من مصادر متنوعة [19]. وستتيح هذه البنية التحتية القوية لمقاييس البيانات المفتوحة رصد الأثر، وتوجيه التمويل المستقبلي، وتحسين نشر البحوث، والمساعدة في إيضاح مشاركة البيانات وإعادة استخدامها ونسبة الفضل فيهما.

الإطار 3. أسئلة العاملين في البحث عن القيم:

- كيف تنسب المؤسسة الفضل إلى الباحثين الذين يتبعون ممارسات (FAIR)؟
- كيف يُعترف بمشاركة البيانات وإعادة استخدامها في قرارات التعيين والترقية؟
- أين يمكنني التعرف على باحثين آخرين يولّدون بيانات متعددة الأنماط والتواصل معهم؟

القاعدة 7: وفّر البنية التحتية لدعم مشاركة البيانات:

يجب أن يقترن الدعم المؤسسي لمشاركة البيانات باستثمار متواصل في تطوير البنية التحتية التقنية والاجتماعية. ومن دون ذلك، سيُنشأ عدد لا يمكن إدارته من الحلول المصممة حسب الطلب للمهام المرتبطة بمشاركة البيانات عبر عمر المشروع البحثي، من بدايته إلى العناية بالبيانات (curation) ونشرها وأرشفتها. ومع أن هذا النهج قد يساعد الباحثين على تلبية الاحتياجات أو المتطلبات الآنية، فإنه غير كفء وغير مستدام معًا. فهو يخلق مزيجًا غير متجانس من الاستراتيجيات والخدمات والأدوات التي يستحيل صيانتها على نطاق واسع. وإضافة إلى ذلك، فإن غياب البنية التحتية يجعل الحوكمة الهادفة (القاعدة 4) أمرًا صعبًا.

ومن أمثلة البنية التحتية أدوات جمع البيانات (مثل السجلات الصحية الإلكترونية)، والتخزين لكلٍّ من بيانات العمل والبيانات الأرشيفية (مثل مستودعات بيانات البحث، والمستودعات التخصصية والعامة)، والنشر (مثل المستودعات المؤسسية، (institutional repositories)). ويمكن لأقسام تقنية المعلومات، ومكاتب النزاهة، والمرافق البحثية الأساسية، والمكتبات، وغيرها، تقديم المشورة بشأن الاعتبارات اللازمة، مثل الأمن والخصوصية، والتشغيل البيني، ومتطلبات بيانات (FAIR). وتشمل البنية التحتية الاجتماعية موارد لدعم التدريب، وتطوير مجتمعات المستخدمين، واستشراف الاحتياجات الناشئة. وفي نهاية المطاف، تتطلب البنية التحتية الناضجة على المستوى المؤسسي تأييدًا من صانعي القرارات المالية والسياساتية، وقبولًا من المستخدمين النهائيين، وحلقات للتغذية الراجعة توجّه التحسين المستمر.

وتختلف قدرة البنية التحتية من مؤسسة إلى أخرى، وسيعكس الاستثمار الأولويات والإمكانات المحلية. وكما ذُكر أعلاه، يمكن للمؤسسات استخدام «مشاركة بيانات البحث: نموذج نضج للقدرة المؤسسية» لحصر الموارد والخدمات والخبرات القائمة والمرغوبة وتقييمها وترتيب أولوياتها. وتتيح البنية التحتية المشتركة والقابلة لإعادة الاستخدام لمشاركة البيانات للمؤسسات الاستفادة من وفورات الحجم والوصول إلى بنية تحتية كانت ستظل لولا ذلك بعيدة المنال أو هامشية مقارنةً بالاحتياجات الأخرى. فعلى سبيل المثال، يمكن للمؤسسات استخدام برمجيات مستودعات

مفتوحة المصدر، مثل (Harvard Dataverse) [20] أو (InvenioRDM) [21] الذي يشغّل (Zenodo) [22]، أو النظر في التعاقد مع مزوّد خدمة مستودع عام، مثل (Dryad) [23]. وتمتد فوائد هذه البنى التحتية المشتركة لمشاركة البيانات وغيرها إلى ما هو أبعد من توفير التكاليف. فالحوكمة والإدارة المجتمعيتان اللتان تقوم عليهما تدفعان الابتكار والامتثال للممارسات الناشئة التي يصعب بغير ذلك تحقيقها وتوفير الموارد لها.

الإطار 4. أسئلة العاملين في البحث عن البنية التحتية:

- ما خيارات تخزين البيانات طويل الأمد المتاحة في المؤسسة، وكم تكلف؟
- هل توجد موارد متاحة يمكن أن تساعدني في إعداد خطة لإدارة البيانات؟
- هل يوجد من يمكنه مساعدتي في إزالة المعرّفات الشخصية من مجموعة بياناتي؟
- كيف يمكنني مشاركة مجموعات بيانات كبيرة مع المتعاونين معي؟

القاعدة 8: استفد من التدريب والموارد لدعم قوى عاملة واسعة التنوع في مجال البيانات:

تتطلب مشاركة البيانات طائفة من الأنشطة والمهارات تتجاوز قيام باحث منفرد برفع مجموعة بيانات إلى مستودع. ويحتاج الأشخاص المنخرطون في مشاركة البيانات إلى الوصول إلى تدريب يدعم أدوارهم المحددة ويراعي طائفة من مستويات المعرفة.

ولحسن الحظ، كانت مجتمعات كثيرة سبّاقة إلى تطوير بنية تحتية تدريبية لمشاركة البيانات ونشرها، يمكن للمؤسسات، بل ينبغي لها، الاستفادة منها.

وعلى المستوى المحلي، كثيرًا ما تقود مكتبات الحرم الجامعي جهود التدريب على مشاركة البيانات، وغالبًا ما تتعاون مع وحدات تقنية المعلومات ومكاتب البحث داخل الجامعة. وتتيح البنية التحتية التدريبية المحلية المتينة لأعضاء القوى العاملة الاستفادة من إرشادات تعكس السياسات والإجراءات والموارد المؤسسية. وعلاوة على ذلك، يمكن لفعاليات التدريب والمنتديات المحلية (مثل التوثيق التعاوني، والقوائم البريدية، ومنصات التواصل) أن تعزز المجتمع والتعاون. وقد استثمرت مؤسسات في تطوير مجموعة واسعة من الموارد على المستويات الوطنية والدولية والتخصصية واستدامتها. ومن الأمثلة البارزة مجتمع (Carpentries)، وهو مجتمع عالمي يعلّم مهارات البيانات والحوسبة لإجراء بحث كفء ومفتوح وقابل لإعادة الإنتاج [24]؛ وبرنامج (DataWorks!) التابع لاتحاد (FASEB)، وهو مبادرة تعزز أفضل الممارسات في مشاركة البيانات وإعادة استخدامها للنهوض بصحة الإنسان [25]؛ والمركز الوطني لخدمات البيانات (National Center for Data Service, NCDS) التابع لشبكة المكتبة الوطنية للطب (Network of the National Library of Medicine, NNLM)، الذي يقدّم تدريبًا لزيادة القدرة في علم البيانات لدى اختصاصيي المعلومات [26]. وتوسّع هذه المبادرات وغيرها توسيعًا كبيرًا قدرة المؤسسة، وقدرتنا الجماعية، على بناء المعرفة في مجال مشاركة البيانات.

الإطار 5. أسئلة العاملين في البحث عن التدريب:

- ما التدريب على مشاركة البيانات الذي تقدّمه المؤسسة؟

- هل يمكنني الحصول على ساعات معتمدة للتعليم المستمر أو شهادات عند المشاركة في التدريب على مشاركة البيانات؟
- هل يوفر قسمي أموالًا للتطوير المهني من أجل التدريب على مشاركة البيانات؟
- هل يمكن للطلاب المشاركة في التدريب على مشاركة البيانات؟

القاعدة 9: التزم بمعايير جودة البيانات وإعادة استخدامها وبأفضل الممارسات:

لتيسير الاكتشاف العلمي وتهيئة مجتمعها البحثي للنجاح، ينبغي للمؤسسات أن تلتزم التزامًا فعليًا بأفضل الممارسات التي تيسّر إعادة الاستخدام وأن تدعمها للنهوض بمشاركة البيانات. وتشمل هذه، على سبيل المثال لا الحصر، مبادئ (FAIR)، التي وجّهت أفضل ممارسات مشاركة البيانات المقروءة آليًا منذ نشرها عام 2016، وطيفًا واسعًا من المعايير التي تحكم بنية البيانات ومحتواها ووصفها [27]. فعلى سبيل المثال، يمكن لباحثي الطب الحيوي استخدام عناصر البيانات المشتركة (common data elements) لتنظيم جمع البيانات وتمثيل تشخيصات الأمراض والأدوية والإجراءات والفحوص المخبرية تمثيلًا متسقًا [28]. ويضمن اتباع أفضل ممارسات البيانات الوصفية، مثل مخطط البيانات الوصفية (DataCite) [29]، قابليةً واسعة لاكتشاف البيانات والبيانات الوصفية عبر أدوات الاكتشاف (discovery) مثل (Google Dataset Search) [30] و(DataCite Commons) [31].

وثمة فرص كثيرة أمام المؤسسات للاستفادة من الحوكمة والبنية التحتية والتدريب لتشجيع الباحثين على تبني أفضل الممارسات واستخدامها. فعلى سبيل المثال، يمكنها إنشاء موارد معلوماتية وصيانتها تحيل إلى المعايير المنطبقة، وتقديم تدريب على دمجها ضمن البنية التحتية البحثية القائمة (مثل دفاتر المختبر الإلكترونية). وينبغي للمؤسسات أن تستثمر في القدرة على العناية بالبيانات، وأن تنظر في إرساء «فحوص عشوائية للممارسات الجيدة في البيانات» لضمان التزام الممارسات الحالية بالإرشادات التي وضعتها المؤسسة وبأفضل ممارسات المجتمع. ويمكن للمؤسسات ذات الاستثمارات الكبيرة في مجالات بحثية محددة أن تنظر في استحداث أدوار لكبار مسؤولي المعلومات (Chief Information Officers) أو كبار مسؤولي البيانات (Chief Data Officers) لتعزيز هياكل الحوكمة القائمة. وأخيرًا، يمكن للمؤسسات أن تشارك مشاركة فعلية في المنظمات التي تدعم تطوير معايير البيانات وتبنيها، ولا سيما في المجالات ذات الصلة بمحفظتها البحثية. وتشمل هذه منظمات تطوير المعايير الخاصة بمجالات بعينها، مثل (Health Level Seven International) [32]، وعلوم البيانات الصحية الرصدية والمعلوماتية (Observational Health Data Sciences and Informatics, OHDSI) [33]، والتحالف العالمي للجينوميات والصحة (Global Alliance for Genomics and Health, GA4GH) [34]، إضافة إلى المجتمعات المكرسة للنهوض بإنشاء مخرجات بحثية قابلة للمشاركة ونشرها، مثل (Mobilizing Computable Biomedical Knowledge) [35].

الإطار 6. أسئلة العاملين في البحث عن جودة البيانات:

- أين يمكنني العثور على موارد تساعدني على فهم جودة البيانات؟
- هل يمكنني الدفع مقابل خدمات استشارية أو خدمات للعناية بالبيانات لمساعدة مختبري في هذا العمل؟
- ما الرخصة التي ينبغي أن أطبقها على بياناتي؟
- هل توجد معايير للحد الأدنى من المعلومات للتجارب التي يجريها مختبري؟

القاعدة 10: أدمج أُطر أخلاقيات البيانات وممارساتها وحسّنها:

ترتبط مشاركة البيانات وأخلاقيات البيانات ارتباطًا لا ينفصم. وينبغي للمؤسسات أن تتعامل عن قصد وباستمرار مع الشواغل والمبادئ الأخلاقية لدعم مشاركة البيانات ومعالجة التحيزات وأوجه الظلم المنهجية. ويعرّف إطار أخلاقيات البيانات (Data Ethics Framework) في الاستراتيجية الفدرالية للبيانات (Federal Data Strategy) أخلاقيات البيانات بأنها «معايير السلوك التي تعزز الأحكام الملائمة والمساءلة عند اقتناء البيانات أو إدارتها أو استخدامها، بهدف حماية الحريات المدنية، وتقليل المخاطر على الأفراد والمجتمع، وتعظيم الصالح العام» [36]. وكما هو الحال مع أفضل ممارسات جودة البيانات وإعادة استخدامها، ينبغي للمؤسسات أن تعالج القضايا الأخلاقية مبكرًا وعلى امتداد عمليات مشاركة البيانات، ويمكنها استخدام الحوكمة والبنية التحتية والتدريب لتحقيق ذلك.

ومن الأطر البالغة الأهمية التي ينبغي النظر فيها مبادئ (CARE) لحوكمة بيانات الشعوب الأصلية (CARE Principles for Indigenous Data Governance) ((Collective benefit): المنفعة الجماعية، وAuthority to control: سلطة التحكم، و(Responsibility): المسؤولية، و(Ethics) : الأخلاقيات) [37]. ومن خلال تفعيل مبادئ (CARE) مع مبادئ (FAIR)، يمكن دعم حقوق السيادة على البيانات وتأكيدها عبر قابلية التنفيذ الآلي، ودمج التركيز على الناس والغاية، واحترام حقوق الشعوب الأصلية في بياناتها ومصالحها فيها [38]. وإضافة إلى ذلك، يقدّم إطار أخلاقيات البيانات في الاستراتيجية الفدرالية للبيانات مجموعة من «مبادئ أخلاقيات البيانات» (Data Ethics Tenets) التي يمكن أن تكون نقطة مرجعية لمعالجة الجوانب المحلية لمشاركة البيانات وتوجيه المساءلة [36].

الإطار 7. أسئلة العاملين في البحث عن الأخلاقيات:

- أين يمكنني معرفة المزيد عن كيفية تطبيق ممارسات أخلاقية في التعامل مع البيانات؟
- إلى جانب الخصوصية، ما الذي أحتاج إلى مراعاته عند إعداد خطتي لمشاركة البيانات؟
- هل لدى المؤسسة موارد قائمة لدعم مبادئ (CARE) وسيادة البيانات؟
- كيف يسهم شركاء البحث المجتمعيون في توجيه ممارسات المؤسسة وسياساتها في مشاركة البيانات؟

أفكار ختامية:

نهدف بهذه القواعد البسيطة العشر إلى تزويد المؤسسات بفهم شامل للمشهد والعوامل المنخرطة في دعم مشاركة البيانات، للمساعدة في توجيه التخطيط الاستراتيجي لتنمية قدرة المؤسسة على مشاركة البيانات. وقد أُعدّ هذا العمل مع مراعاة عدة اعتبارات. أولًا، القواعد والخصائص البيئية التي تصفها مترابطة. فمن الصعب الارتقاء بالنضج في مجال ما دون معالجة نقاط القوة والضعف في مجال آخر. ثانيًا، نموذج نضج القدرة على مشاركة بيانات البحث ومستويات النضج التي يحددها يعبران عن طموحات وأهداف تسعى المؤسسات إليها في تقدمها. وكما ذُكر أعلاه، تقف المؤسسات البحثية عند مراحل مختلفة من رحلتها المحلية، ويساعد هذا الإطار في التخطيط لمبادرات التحسين.

وفي الوقت الراهن، لا تزال مؤسسات كثيرة في المراحل الأولى من النظر في الخدمات والموارد والبنية التحتية التي ينبغي لها تنفيذها لدعم مشاركة البيانات. ونهدف إلى مساعدة المؤسسات على تنمية نهج متماسك لمشاركة البيانات، وتشكيل تطوير السياسات والبنية التحتية والقيم المؤسسية التي تدعم مشاركة البيانات من أجل

الصالح العام، لا من أجل الامتثال فحسب. ولكل عضو في المؤسسة دور يؤديه في هذه النقاشات. وتقدّم هذه الورقة و«مشاركة بيانات البحث: نموذج نضج للقدرة المؤسسية» إطارًا لتنسيق التواصل والعمل العمليين عبر الأدوار والمسؤوليات على المستويات الفردية والجماعية والمؤسسية.

ويعمل مؤلفو هذه الورقة عادةً في بيئة البحث الطبي الحيوي، التي تتطلب اعتبارات محددة للأمن والخصوصية والأخلاقيات لاستيفاء متطلبات مشاركة البيانات في البحوث على المشاركين من البشر. ومع أننا نشير إلى جوانب من هذا السياق في هذه الورقة ونذكرها أيضًا في نموذج النضج الخاص بنا، فإننا لم نتناول تفاصيل هذا السياق الطبي الحيوي، على أمل أن توسّع مجتمعات التخصصات الأخرى التوصيات والنموذج اللذين طورناهما وتحسّنهما، لتعكس القضايا المحددة، والمختلفة أحيانًا، التي يتعين عليها معالجتها لدعم مشاركة البيانات. ومن ثَمّ ندعو مجتمع مشاركة البيانات الأوسع إلى مراجعة «مشاركة بيانات البحث: نموذج نضج للقدرة المؤسسية» ونقده وتكييفه.

المراجع:

1. NOT-OD-21-013: Final NIH Policy for Data Management and Sharing [Internet]. [cited 2022 Oct 16]. Available from: https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html.
2. Gonzales S, Carson MB, Holmes K. Ten simple rules for maximizing the recommendations of the NIH data management and sharing plan. PLoS Comput Biol. 2022 Aug 3;18(8):e1010397. pmid:35921268
3. Michener WK. Ten Simple Rules for Creating a Good Data Management Plan. PLoS Comput Biol. 2015 Oct 22;11(10):e1004525. pmid:26492633
4. Champieux RE, Phuong J, Dorr D, Holmes K, Rojevsky S, Solomonides A, et al. Research Data Sharing: A Maturity Model for Organizational Capacity. 2022 Nov 27 [cited 2022 Nov 27]. Available from: https://zenodo.org/record/7369811.
5. Crowston K, Qin J. A capability maturity model for scientific data management: Evidence from the literature. Proc Am Soc Inf Sci Technol. 2011;48(1):1-9.
6. Knosp BM, Barnett WK, Anderson NR, Embi PJ. Research IT maturity models for academic health centers: Early development and initial evaluation. J Clin Transl Sci. 2018 Oct;2(5): 289-94. pmid:30828469
7. CTS Personas [Internet]. CTS Personas. 2023 [cited 2023 Mar 22]. Available from: https://zenodo.org/communities/cts-personas.
8. User story. In: Wikipedia [Internet]. 2023 [cited 2023 Mar 23]. Available from: https://en.wikipedia.org/w/index.php?title=User_story&oldid=1145675576.
9. Inclusive Team Science Toolkit-CREDITS [Internet]. [cited 2023 Mar 22]. Available from: https://credits.ucsb.edu/inclusive-team-science-toolkit/.

10. Spring B, Klyachko E, Rak P, McFadden H, Hedeker D, Siddique J, et al. Online, cross-disciplinary team science training for health and medical professionals: Evaluation of COALESCE (teamscience.net). J Clin Transl Sci. 2019 Jul 18;3(2–3):82–9. pmid:31660230

11. Lerner D, Palm ME, Concannon TW, editors. Broadly Engaged Team Science in Clinical and Translational Research [Internet]. Cham: Springer International Publishing; 2022 [cited 2023 Mar 22]. Available from: https://link.springer.com/10.1007/978-3-030-83028-1.

12. Home | Research Data Management Service Group [Internet]. [cited 2023 Mar 22]. Available from: https://data.research.cornell.edu/.

13. Research Data Management @Harvard [Internet]. [cited 2023 Mar 22]. Available from: https://researchdatamanagement.harvard.edu/home.

14. Jenkins D. Introduction to Data Sharing and Integration [Internet]. 2020 [cited 2022 Oct 16]. Available from: https://aisp.upenn.edu/aisp-intro/.

15. National Academies of Sciences E. Open Science by Design: Realizing a Vision for 21st Century Research [Internet]. 2018 [cited 2022 Oct 16]. Available from: https://nap.nationalacademies.org/catalog/25116/open-science-by-design-realizing-a-vision-for-21st-century.

16. Tenopir C, Palmer CL, Metzer L, van der Hoeven J, Malone J. Sharing data: Practices, barriers, and incentives. Proc Am Soc Inf Sci Technol. 2011;48(1):1–4.

17. Piwowar HA, Becich MJ, Bilofsky H, Crowley RS. Towards a Data Sharing Culture: Recommendations for Leadership from Academic Health Centers. PLoS Med. 2008 Sep; 5(9):e183. pmid:18767901

18. Wood-Charlson EM, Crockett Z, Erdmann C, Arkin AP, Robinson CB. Ten simple rules for getting and giving credit for data. PLoS Comput Biol. 2022 Sep 29;18(9):e1010476. pmid: 36173960

19. Vierkant P. Wellcome Trust and the Chan Zuckerberg Initiative Partner with DataCite to Build the Open Global Data Citation Corpus [Internet]. DataCite Blog. 2023 [cited 2023 Mar 22]. Available from: https://blog.datacite.org/data-citation-corpus-announcement-2023/.

20. Harvard Dataverse [Internet]. [cited 2022 Oct 22]. Available from: https://dataverse.harvard.edu/.

21. InvenioRDM—inveniosoftware.org [Internet]. [cited 2022 Oct 22]. Available from: https://inveniosoftware.org/products/rdm/.

22. Zenodo—Research. Shared. [Internet]. [cited 2022 Oct 22]. Available from: https://zenodo.org/.

23. Dryad Our Mission [Internet]. [cited 2022 Oct 22]. Available from: https://datadryad.org/stash/our_mission.

24. The Carpentries [Internet]. The Carpentries. [cited 2022 Oct 16]. Available from: https://carpentries.org/index.html.

25. DataWorks! [Internet]. [cited 2022 Oct 16]. Available from: https://www.faseb.org/resources/data-science-and-informatics/dataworks.

26. NNLM National Center for Data Services (NCDS) [Internet]. [cited 2022 Oct 16]. Available from: https://www.nnlm.gov/about/centers/ncds.

27. Wilkinson MD, Dumontier M, Aalbersberg IjJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016 Mar15;3(1):160018. pmid:26978244

28. NIH Common Data Elements (CDE) Repository [Internet]. [cited 2022 Oct 22]. Available from: https://cde.nlm.nih.gov/home.

29. DataCite Schema [Internet]. DataCite Schema. [cited 2022 Oct 22]. Available from: http://schema.datacite.org/.

30. Dataset Search [Internet]. [cited 2023 Mar 22]. Available from: https://datasetsearch.research.google.com/.

31. DataCite Search [Internet]. [cited 2022 Oct 22]. Available from: https://search.datacite.org/.

32. Health Level Seven International—Homepage | HL7 International [Internet]. [cited 2022 Oct 16]. Available from: https://www.hl7.org/.

33. OHDSI-Observational Health Data Sciences and Informatics [Internet]. [cited 2022 Oct 16]. Available from: https://www.ohdsi.org/.

34. Global Alliance for Genomics and Health (GA4GH) [Internet]. [cited 2022 Nov 27]. Available from: https://www.ga4gh.org/.

35. Mobilizing Computable Biomedical Knowledge [Internet]. [cited 2022 Oct 16]. Available from: https://mobilizecbk.med.umich.edu/.

36. Data Ethics Framework [Internet]. [cited 2022 Oct 16]. Available from: https://strategy.data.gov/assets/docs/2020-federal-data-strategy-framework.pdf.

37. Carroll SR, Garba I, Figueroa-Rodríguez OL, Holbrook J, Lovett R, Materechera S, et al. The CARE Principles for Indigenous Data Governance. Data Sci J. 2020 Nov 4;19(1):43.

38. Carroll SR, Herczog E, Hudson M, Russell K, Stall S. Operationalizing the CARE and FAIR Principles for Indigenous data futures. Sci Data. 2021 Apr 16;8(1):108. pmid:33863927

39. Table 1: https://doi.org/10.1371/journal.pcbi.1011136.t001.