

خدمة واحدة لبيانات البحث: كيف تجمع الجامعات البريطانية مواردها؟

جون كاي، رايتشل بروس، دوم فريب



الملخص:

يتناول المقال، بوصفه دراسة حالة، تطوير (Jisc) لخدمة بيانات البحث المشتركة (RDSS) بالشراكة مع المجتمع البحثي في المملكة المتحدة، لتمكين المؤسسات من دعم الباحثين في إيداع البيانات ونشرها واكتشافها وتخزينها وأرشفتها وحفظها على المدى الطويل. ويعرض المؤلفون دوافع المشروع في مشهد مجزأ لإدارة بيانات البحث، ومراحل جمع المتطلبات من خلال استبانة مؤسسية وبحث مكتبي وورش تشاورية، واختيار المؤسسات التجريبية ومناقصة الموردين بثماني حزم. ويعرضون نتائج استبانة إطار أصول البيانات (DAF) لعام 2016، ونتائج مجموعات التركيز بشأن البيانات الوصفية، وتصور الخدمة، والبنية التقنية القائمة على الأحداث، ونموذج البيانات. ويختتمون بوصف التقدم المحرز في مرحلة التطوير الأولى (ألفا) والتحديات القائمة، كتحديد الحد الأدنى من المنتج القابل للاستخدام، والملاءمة مع مسارات العمل القائمة، وحفظ بيانات البحث، وإدارة البيانات الضخمة والحساسة.

الكلمات المفتاحية: إدارة بيانات البحث، الخدمات المشتركة، الحفظ الرقمي، البيانات الوصفية، (Jisc)

المقدمة والخلفية:

نستخدم في هذا المقال مصطلح «بيانات البحث» وفق تعريف مجلس بحوث الهندسة والعلوم الفيزيائية (Engineering and Physical Sciences Research Council, EPSRC): «تُعرّف بيانات البحث بأنها المواد الواقعية المسجلة التي يحتفظ بها المجتمع العلمي عادةً ويقبلها بوصفها ضرورية للتحقق من نتائج البحث؛ ومع أن معظم هذه البيانات يُنشأ بصيغة رقمية، فإن جميع بيانات البحث مشمولة بصرف النظر عن الصيغة التي أنشئت بها».[1]

منذ عدد من السنوات، وإدارة بيانات البحث مدرجة على جدول أعمال الجهات الممولة للبحوث والمؤسسات البحثية، وقد عملت (Jisc) مع الجامعات والجهات الممولة سعيًا إلى تلبية الاحتياجات المتصلة بها. وقرب نهاية برنامج (Jisc) لإدارة بيانات البحث، لخص مشروع (Data Pool) في جامعة ساوثهامبتون (University of Southampton) الوضع على النحو الآتي:

«... من الواضح أن القضايا المحيطة بإدارة بيانات البحث تزداد تعقيدًا لا بساطة. فنحن نفهم الآن أكثر بكثير عن نطاق البيانات التي يتعين إدارتها، وحجمها وتعقيدها، وتوقعات الباحثين في إدارة مسارات العمل ومشاركة البيانات. ونعلم أيضًا أن متطلبات الحكومة والجهات الممولة تفرض على المؤسسات، على المستوى المؤسسي، تكاليف مالية قد تكون كبيرة، وتجد المؤسسات صعوبة في الوفاء بها في المناخ المالي الحالي».[2]

وهذا يصف البيئة التي برزت فيها خدمة بيانات البحث المشتركة (RDSS) [3] أولويةً تطويرية لدى (Jisc). فمن خلال العملية التشاورية المحيطة بتحدي التصميم المشترك «البحث في خطر» (Research at Risk) [4]، الذي قاده (Jisc) بالشراكة مع رابطة المكتبات البحثية في المملكة المتحدة (RLUK)، ومنتدى مديري تقنية المعلومات في مجموعة جامعات راسل (RUGIT)، وجمعية مكتبات الكليات والمكتبات الوطنية والجامعية (SCONUL)، ورابطة نظم المعلومات في الجامعات والكليات (UCISA)، جرى تحديد الأهداف العليا الآتية التي ستحتاج مؤسسات التعليم العالي (HEIs) إلى خدمة مشتركة لتحقيقها:

- الامتثال لسياسات بيانات البحث، مثل الوفاء باشتراطات (EPSRC) [5]
- زيادة الكفاءة في القطاع، مثل الشراء المشترك، والاستخدام المشترك للنظم، وفرص إعادة استخدام البيانات، وفرص التشغيل البيني
- تحسين نزاهة البحث بإتاحة البيانات لإعادة إنتاج نتائج البحوث
- نظام متكامل لإدارة بيانات البحث (research data management, RDM) من البداية إلى النهاية، يسد فجوة الحفظ الرقمي، ويحسن تجربة المستخدم وقابلية الاستخدام
- تسريع إدارة بيانات البحث في المؤسسات
- توفير حل قادر على استيعاب مخرجات البحث من البيانات والنصوص، ومن ثمّ دعم المؤسسات في الوفاء بمتطلبات الوصول المفتوح وإطار التميز البحثي (Research Excellence Framework, REF) [6].

في البداية، كانت الحاجة إلى معالجة النهج غير الكفاء في إدارة بيانات البحث هي محرك المشروع، ومن خلال التشاور مع القطاع برزت صورة لمشهد مجزأ لإدارة بيانات البحث. فقد تألف هذا المشهد من عدد من الخدمات التجارية والمفتوحة المصدر القائمة التي تلبى متطلبات أجزاء من الخدمات اللازمة لتحقيق أهداف الجامعات، لكن لم يكن هناك نظام واحد يلبيها جميعًا. وقد يكون شراء جميع المكونات اللازمة لحل متكامل لإدارة بيانات البحث وتطويرها وتجميعها مستنزفًا لموارد المؤسسات. ويمكن أن تؤدي هذه السوق وتباين مناهج المؤسسات إلى نظم فيها فجوات وظيفية، أو إلى مؤسسات بلا نظم لإدارة بيانات البحث على الإطلاق. ولوحظت فجوات وظيفية واسعة الانتشار فيما يتعلق بالتشغيل البيني بين نظم البحث، وغياب إجراءات الحفظ ونظمه فيما يتجاوز النسخ الاحتياطي للبيانات. ومن شأن نظام تشتره (Jisc) وتديره وتستضيفه أن يرفع عن المؤسسات عبء الشراء والإدارة والتطوير هذا، وأن يحقق وفورات مالية في القطاع، إلى جانب تمكين أفضل الممارسات في إدارة بيانات البحث والوفاء باشتراطات الجهات الممولة.

وفي الوقت نفسه الذي بدأ فيه عملنا الأولي في أواخر 2014 ومطلع 2015، كان هناك نقاش وطني بشأن السياسات حول وضع ميثاق (Concordat) بين الجهات الممولة والجامعات بشأن بيانات البحث المفتوحة. [7] وناقش قادة الجامعات ذلك عبر رابطة جامعات المملكة المتحدة (Universities UK) [8]، وجرى البحث عن حل عملي يدعم تطلعات الميثاق. ولذلك عُدّ مشروع (Jisc) لخدمة بيانات البحث المشتركة تطورًا مهمًا لتلبية هذه الحاجة.

واستجابةً للأولويات العليا لدى المؤسسات، حددت (Jisc) نطاقًا للمشروع الأولي للخدمة. فسنركز على نظام يتيح الاستيعاب (ingest) والنشر والتخزين طويل المدى والحفظ لكائنات البيانات النهائية المعدة للنشر أو الأرشفة، وينشئ روابط بالخدمات القائمة في جزأي «إنشاء البيانات» و«إدارة البيانات النشطة» من دورة الحياة. ومن ثمّ لن يشمل توفير نظم تخزين نشطة «حية» للكائنات التي ينشئها الباحثون ويعملون عليها ضمن مسارات عملهم الخاصة.

جمع المتطلبات:

لإنتاج نظام يلائم أكبر قدر ممكن من قطاع التعليم العالي، تُقَدِّم عملية شاملة لجمع المتطلبات في النصف الثاني من عام 2015. وتألّفت هذه العملية من ثلاثة مكونات رئيسية:

- استبانة مؤسسية حول نظم البحث
- بحث مكتبي حول المتطلبات القائمة
- ورش عمل لجمع المتطلبات.

صُمِّمَت الاستبانة المؤسسية حول نظم البحث للنظر في الاستخدام الحالي لنظم البحث، والحالة الراهنة للجاهزية في إدارة بيانات البحث، واحتياجات تخزين البيانات الحالية والمستقبلية، وما يريد القطاع من (Jisc) أن تقدمه في مجال إدارة بيانات البحث. وأظهرت الإجابات تفاوتًا في درجة النضج، إذ كانت لدى بعض الجامعات خطط قائمة بالفعل، لكن نسبة كبيرة من المجيبين طلبت تدخلًا من (Jisc)، وكان واضحًا أن احتياجات بيانات البحث ستزداد زيادة كبيرة خلال السنوات الثلاث التالية.

ونظر البحث المكتبي في المتطلبات القائمة لنظم إدارة بيانات البحث التي وضعتها المؤسسات ضمن برنامج عمل (Jisc) السابق «إدارة بيانات البحث» [9]، وفي المتطلبات المنبثقة عن مبادرة (Jisc) «ربيع بيانات البحث» (research data spring) [10]. ومن المشروعات التي استند إليها في البداية:

- خريطة طريق المشروع التجريبي لإدارة بيانات البحث في جامعة ليدز (University of Leeds) [11]
- مسودة مواصفات المتطلبات في جامعة مانشستر متروبوليتان (Manchester Metropolitan University)
- مشروع (ADMIRE) [12]
- (Kaptur) [13]
- (SWORD) [14]
- سدّ فجوة حفظ البيانات (Filling the Data Preservation Gap) [15]
- نهج اتحادي لنظام متكامل لإدارة بيانات البحث (A consortial approach to integrated RDMS) [16]

• (Data Vault) [17]

• المتطلبات التشغيلية لتحليلات المتعلمين لدى (Jisc). [18]

وَدُمجت هذه المتطلبات مع نتائج الاستبانة وُجِّمَت في وثيقة متطلبات شاملة [19] تشاور المجتمع بشأنها ورَّب أولوياتها خارج الورش وفي ورش تشاورية. ولم نرد أن ننشئ متطلبات جديدة من الصفر، بل أن نستخدم أفضل الممارسات من المملكة المتحدة ومن أنحاء العالم.

واتخذت الورش التشاورية شكل ورشة صغيرة [20] لجسّ النبض بأفكار مفاهيمية، واستكشاف بنى محتملة (مثل تلك التي أعدتها جامعة شيفيلد (University of Sheffield)) [21]، والحصول على مدخلات خبيرة من مؤسسات كانت تتعامل بالفعل مع نظم إدارة بيانات البحث، ثم نُظِّمَت فعالية أكبر [22] شارك فيها 70 ممثلًا من مؤسسات التعليم العالي في المملكة المتحدة، وقُدِّمَت آراء بشأن الأولويات والمتطلبات، وتبادل المشاركون أيضًا المعارف والخبرات بشأن الفجوات الحالية في توفير خدمات إدارة بيانات البحث، والدروس المستفادة من تطبيق الخدمات ذات الصلة على المستوى المؤسسي، وناقشوا التطلعات والمخاوف بشأن الخدمات المشتركة لإدارة بيانات البحث.

وكان هناك توافق على أن تقوم (Jisc) بما يأتي:

• إنتاج نظام جديد يمكن تقديمه خدمة مُدارة، فيرفع العبء عن موظفي تقنية المعلومات والمشتريات في المؤسسات

• تطوير تشغيل بيني محسّن، على نحو جماعي، بين نظم الخدمة والنظم المؤسسية والخارجية القائمة

• إنتاج نظام متكامل من البداية إلى النهاية، مع إتاحة خيارات «اختيار المكونات ودمجها» لتلبية جميع احتياجات الجامعات

• شراء خدمات واستشارات في إدارة بيانات البحث لدعم متطلبات المؤسسات التجريبية في هذا المجال وتطبيقها.

وعُقدت أيضًا ورشة أصغر مع موردي النظم المحتملين، قدّموا فيها آراءهم ونصائحهم بشأن المقترحات والمتطلبات التي أعدها المجتمع. وتمثلت النتيجة النهائية لهذه المرحلة من جمع المتطلبات وتحليلها في المتطلبات التشغيلية للخدمة التي نُشرت ضمن مناقصة الخدمة في الجريدة الرسمية للاتحاد الأوروبي (The Official Journal of the European Union, OJEU). [23]

بناء الفريق: المؤسسات التجريبية:

لما كان توجيه القطاع أن أي حل يجب أن يلبي طيفًا من الاحتياجات المؤسسية، سعت (Jisc) إلى الشراكة مع عدد من مؤسسات التعليم العالي التجريبية المتنوعة لتطوير الخدمة. ولإنشاء «مجموعة متوازنة»، قُيِّمَت الجامعات التي أبدت اهتمامًا بالانضمام إلى المرحلة التجريبية وفق معايير تراعي حجم المؤسسة ونوعها (مثل المؤسسات كثيفة البحث، والصغيرة والمتخصصة، والقائمة على التدريس)، وتوافر أنواع متباينة من البيانات، ودرجة جاهزيتها في إدارة بيانات البحث (من المواقع التي تبدأ من الصفر إلى الأكثر رسوخًا)، والاستخدام الحالي للنظم المؤسسية. [24]

ونتيجة لعملية الاختيار هذه، تتعاون (Jisc) مع المؤسسات التجريبية الآتية لتطوير الخدمة:

• جامعة كارديف (Cardiff University)

• (CREST) (اتحاد التميز البحثي والدعم والتدريب، Consortium for Research Excellence, Support and Training)

• إمبريال كوليدج لندن (Imperial College London)

• جامعة بليموث (Plymouth University)

• جامعة ميدلسكس (Middlesex University)

• الكلية الملكية للموسيقى (Royal College of Music)

• كلية سانت جورج الطبية (St George's Hospital Medical School)

• جامعة كامبريدج (University of Cambridge)

• جامعة لانكستر (Lancaster University)

• جامعة لينكولن (University of Lincoln)

• جامعة سانت أندروز (University of St Andrews)

• جامعة سري (University of Surrey)

• جامعة يورك (University of York).

وقد رُتبت أولويات طائفة من الأهداف التي طرحتها المؤسسات التجريبية، ومن ملخص انتقائي لها:

• حل سهل الاستخدام ومنخفض التكلفة للأرشفة والاستيعاب والحفظ والمستودع وإعداد التقارير والاكتشاف، قادر على التعامل مع البيانات الحساسة

• تخزين متين للبيانات قابل للتوسع للبيانات النشطة والمؤرشفة

• ملف تعريف معياري للبيانات الوصفية، دولي لتحقيق التشغيل البيئي

• التكامل مع جميع نظم معلومات البحث الجارية (Current Research Information Systems, CRIS) الرئيسة

• الوفاء بمتطلبات الإيداع لإطار التميز البحثي والجهات الممولة (دعم إيداع أنواع مخرجات البيانات في إطار التميز البحثي).

وجرى جمع المتطلبات التقنية ومتطلبات المستخدمين بتعمق مع كل مؤسسة تجريبية، وجمّعت هذه المتطلبات ورُتبت أولوياتها على مستوى المؤسسات كلها لتحديد مسار التطوير.

بناء الفريق: الموردون:

مع أن بعض التطوير المخصص سيكون مطلوبًا، وأن هناك قيمة مضافة ينبغي تطويرها على امتداد النظام المتكامل، فقد تقرر البناء على ما هو موجود بالفعل. ولذلك استخدمت (Jisc) إجراءات الشراء وفق الجريدة الرسمية للاتحاد الأوروبي لإنشاء إطار للموردين يُختار منه موردو النظم ومطوروها لإنشاء الخدمة. وقُسم الإطار إلى ثماني حزم على النحو الآتي:

• الحزمة 1: موردو مستودعات بيانات البحث

- الحزمة 2: موردو واجهات المستودعات
- الحزمة 3: موردو واجهات تبادل بيانات البحث
- الحزمة 4: موردو تكامل نظم معلومات البحث وإدارته
- الحزمة 5: موردو منصات حفظ بيانات البحث
- الحزمة 6: موردو تطوير أدوات حفظ بيانات البحث
- الحزمة 7: موردو تقارير بيانات البحث
- الحزمة 8: موردو تحسين تجربة المستخدم.

ويمكن الاطلاع على مزيد من المعلومات عن تفاصيل وظائف هذه الحزم في وثيقة المتطلبات التشغيلية للخدمة [25]، وعلى قائمة الموردين الفائزين بالحزم في مدونة (Jisc) لإدارة بيانات البحث. [26]

تلبية احتياجات الباحثين: إطار أصول البيانات:

ستتيح الخدمة للباحثين إيداع البيانات لنشرها واكتشافها وتخزينها الآمن وأرشفتها على المدى الطويل وحفظها. غير أن ذلك يشير طائفة كاملة من الأسئلة المختلفة، منها:

- ما أشكال البيانات التي لدى الباحثين؟
- ما حجم البيانات الذي نتحدث عنه؟
- أين يخزن الباحثون بياناتهم حالياً؟
- من يحتاج أيضاً إلى الوصول إليها؟
- إلى متى يجب الاحتفاظ بالبيانات؟
- ما الذي يحفز الباحثين على مشاركة بياناتهم، أو على إبقائها مغلقة؟

وعلى مدار عام 2016، عملت (Jisc) مع (Research Consulting) [27] ومع المؤسسات التجريبية لمحاولة إيجاد بعض الإجابات.

ولتنفيذ هذا العمل، تقرر أن إطار أصول البيانات (Data Asset Framework, DAF) [28] يقدم منهجية مفيدة جربتها مؤسسات في أنحاء العالم واختبرتها. وقد طوّر هذا الإطار عام 2009 لمساعدة المؤسسات على تحديد أصول بيانات البحث لديها وتعيين مواقعها ووصفها وتقييم طريقة إدارتها. ويستخدم الاستبانة والمقابلات لجمع المعلومات اللازمة، واخترنا اتباع نهج مماثل. والأفضل من ذلك أن ست مؤسسات من المؤسسات التجريبية للخدمة كانت قد أجرت بالفعل استبانة الإطار خلال العامين السابقين، فوقّر لنا ذلك بعض البيانات الجاهزة. غير أننا حين شرعنا في تحليل هذه البيانات وجدنا أن كل مؤسسة عدّلت أسئلة الاستبانة إلى حد ما، فبقيت لدينا مجموعة من النتائج الصالحة كلّ على حدة، لكن لا يمكن تجميعها تجميعاً ذا معنى. يضاف إلى ذلك أن مشهد بيانات البحث تطور سريعاً منذ عام 2009. فسياسات الجهات الممولة بشأن مشاركة البيانات صارت أشد صرامة بكثير، واستخدام نظم

معلومات البحث الجارية صار أوسع انتشارًا، وظهرت خدمات جديدة مثل (ORCID) [29] و(DataCite) [30]. ونتيجة لذلك، لم تكن مجموعة الأسئلة الأصلية لإطار أصول البيانات [31] تغطي كثيرًا من المعلومات التي كنا نحتاج إلى معرفتها.

وكان الحل البديهي تطوير نسخة جديدة من الاستبانة: «استبانة (DAF) لعام 2016». فانطلاقًا من نسخة عام 2009، أعدّ موظفو (Jisc) والمؤسسات التجريبية للخدمة مجموعة أسئلة منقحة بين أبريل ويونيو 2016. ثم أجرت ست مؤسسات تجريبية الاستبانة الجديدة في يوليو وأغسطس من ذلك العام، وشملت البيانات 1,185 إجابة فريدة. ونُشر ملخص لأبرز النتائج في عرض شرائح. [32]

وتتيح لنا بيانات الاستبانة استخلاص بعض الاستنتاجات الأعم بشأن الحالة الراهنة لإدارة بيانات البحث في المملكة المتحدة:

- يمكن للخدمة أن تسد فجوة مهمة: فـ75% من الباحثين يتطلعون أولًا إلى مؤسستهم لحفظ بياناتهم، لكننا نعلم أن مؤسسات كثيرة لا تستطيع حاليًا تلبية هذه الحاجة تلبية كاملة. وهنا يمكن للخدمة أن تساعد.
- لا يزال الإقبال على الدعم المؤسسي لإدارة بيانات البحث منخفضًا: فـ16% فقط من المجبيين يستفيدون حاليًا من خدمات الجامعة لدعم إدارة بيانات البحث. وهذا تحدٍّ مزدوج: فعلى المؤسسات ألا تكتفي بإتاحة خدمات الدعم المناسبة، بل أن تعرّف الباحثين بوجودها أيضًا.
- هناك استعداد قائم للمشاركة: فـ68% من المجبيين إما يشاركون البيانات بالفعل وإما يتوقعون أن يفعلوا ذلك في المستقبل. ويفعل معظمهم ذلك لأنهم يؤمنون بأن البحث منفعة عامة ينبغي أن تكون مفتوحة للجميع. وكل ما نحتاج إليه هو جعل مشاركة البيانات أيسر.
- لا يزال أماننا طريق طويل: فـ40% فقط من المجبيين لديهم حاليًا خطة لإدارة بيانات البحث، و18% فقط يتبعون معايير أو إرشادات راسخة للبيانات الوصفية. وإحداث التغيير سيستغرق وقتًا.

ويتوافر تقرير كامل عن نتائج الاستبانة [33]، إلى جانب مجموعة البيانات الخام المجهّلة الداعمة له. [34]

وإلى جانب الاستفادة من هذا العمل في تحديد متطلبات الخدمة، أعدنا موارد يمكن إعادة استخدامها. فقد نُشرت حقيبة أدوات جديدة لإطار أصول البيانات تبين الخطوات اللازمة لإجراء استبانة الإطار، وتقدم توصيات بشأن كيفية تناولها.

تلبية احتياجات الباحثين: البيانات الوصفية:

لما كانت الخدمة تهدف إلى إنشاء خدمة بيانات متعددة التخصصات، كان علينا التأكد من أن نماذج البيانات الوصفية ونماذج البيانات لدينا والعمليات المرتبطة بها قادرة على دعم حالات الاستخدام لدى طيف واسع من الباحثين من مجالات بحثية شديدة التنوع، ولذلك احتجنا إلى اختبار أفكارنا مع الباحثين. ولتحقيق ذلك، عملت (Jisc) مع (Clax) [35] لعقد تسع مجموعات تركيز مع باحثين من المؤسسات التجريبية. ويتوافر تقرير كامل عن مجموعات التركيز [36] إلى جانب مجموعة شاملة من حالات استخدام الباحثين [37] ونعرض فيما يأتي بعض المحاور الناشئة:

- أبدت مجموعات التركيز قلقها بشأن عدد من الجوانب المتعلقة بالبيانات الوصفية. ويمكن معالجة بعضها بالتدريب والدعم، ويمكن معالجة كثير منها بعمل الموردين مع المؤسسات ومع الخدمة. وقليل منها يتطلب تقنيات جديدة أو تغييرًا ثقافيًا.
- كثيرًا ما ذكر الإنشاء والجمع المبكر للبيانات الوصفية. ويمكن تحقيق ذلك باستخدام خطط ديناميكية لإدارة البيانات، بحيث تُجمع البيانات الوصفية منذ مرحلة التخطيط وتُحدَّث طوال عملية جمع البيانات وتحليلها.
- ينبغي للنظم أن تحفظ شكل البيانات المودعة ومحتواها، مع إتاحة تحديث البيانات الوصفية لربطها بمجموعات البيانات ذات الصلة، والمنشورات اللاحقة، والمواد الأخرى التي ربما أُنشئت بعد إيداع البيانات. وينبغي لها أيضًا أن تتيح تحديث الكلمات المفتاحية والمواد الوصفية لتعكس التغيرات في التخصص. ومن الضروري توافر إمكان تضمين البيانات الوصفية روابط إلى معرفات الكائنات الرقمية (DOIs) الأخرى، وإلى عناوين (URL) حين لا يوجد معرف (DOI).
- كثيرًا ما يُفترض أن جمع البيانات الوصفية سيُلزم الباحثين بملء نماذج شاقة ومستهلكة للوقت عند إيداع البيانات. وهذا غير مرغوب فيه، ولا يُرجَّح أن ينتج بيانات وصفية جيدة. وبدلًا من ذلك، فإن أتمتة الأدوات وعمليات الجمع والمعدات، وجمع البيانات الوصفية على نحو مدمج في مسار عمل الباحثين طوال البحث، سيتيحان، في الوضع المثالي، إرسال البيانات إلى المستودع بضغط زر، مع إرفاق البيانات الوصفية بها مسبقًا.

طموحات الخدمة:

أدى جمع المتطلبات وتحديد احتياجات المستخدمين إلى رسم تصور للخدمة، تقوم فيه المنصات الأساسية على نظم المستودع والحفظ وإعداد التقارير، مدعومةً بتخزين للبيانات توفره (Jisc). [38] ويشمل ذلك المدخلات من الباحثين عبر واجهة مستخدم على الويب وأدوات أخرى لإدارة البيانات الضخمة والحساسة. والتشغيل البيئي عنصر أساسي في الخدمة، وقد حُدِّد عدد من حالات التكامل، مثل التكامل مع نظم (CRIS)، وأدوات تخطيط إدارة البيانات، ومستودعات المنشورات. ولن تعمل الخدمة بمعزل عن غيرها؛ إذ ستُنشأ روابط مع البنية التحتية للاتصال العلمي من خلال استخدام المعرفات الدائمة، مثل معرفات (DOI) ومعرفات (ORCID)، وخدمات التجميع، وستستخدم الخدمة البنية التحتية للمقاييس والمقاييس البديلة (altmetrics)، وستتواصل مع نظم الجهات الممولة حيثما أمكن.

وستتربط الخدمة أيضًا مع خدمات وسياسات وممارسات ومعايير أخرى يجري تطويرها ضمن مجموعة مشروعات (Jisc) «البحث في خطر»، مثل التشغيل البيئي مع خدمة اكتشاف بيانات البحث في المملكة المتحدة (UK Research Data Discovery Service) [39] وتزويدها بالبيانات الوصفية، ودمج أعمال المقاييس وإحصاءات

الاستخدام المطوّرة في (IRUS Data UK) [40]، ودمج (ORCID) [41] في الخدمة. وقد استرشدت متطلبات المرحلة التجريبية بالفعل بإرشادات (Jisc) بشأن سياسات الجهات الممولة [42]، وستسعى إلى تسخير الابتكار الناتج عن المشروعات الناجحة في «ربيع بيانات البحث».

البنية التقنية ونهج التطوير:

تناول معظم هذا المقال إرساء الأسس لتقديم الخدمة. غير أن (Jisc) دخلت مرحلة الإنتاج التقني للنسخة الأولى (ألفا) من الخدمة في نوفمبر 2016. وجرى الاتفاق على نهج للبنية التقنية مع المؤسسات التجريبية والموردين والفريق الاستشاري الخبير في أكتوبر 2016. ومن العناصر الرئيسة لهذا النهج:

- الاتفاق على أن البنية الموجهة نحو الخدمات والقائمة على الأحداث (SOA 2.0) ضرورية، بسبب الحاجة إلى التوسع من حيث التعقيد، وإلى بنية أكثر تحملاً للأعطال. وتوفر (SOA 2.0) أعلى مستوى من الفصل بين الخدمات. وكثيراً ما يصف المهندسون المحترفون نظم (SOA 2.0) بعبارة «قنوات نقل بسيطة ونقاط نهاية ذكية». فبدلاً من تنسيق إجراءات العمل مركزيًا، تُصفي الخدمات إلى الأحداث أو الإشعارات الصادرة عن المكونات الأخرى وتستجيب لها على النحو المناسب. وتتسم (SOA 2.0) بطبيعتها الموزعة.
- تعني هذه البنية أننا نحتاج إلى مجموعة من الخدمات المصغرة (microservices) والمحولات لتحويل البيانات من النظم الخارجية إلى نموذج البيانات المعياري (canonical data model) لدينا.
- نمط تكامل النظم هو النشر والاشتراك (publish-subscribe).

ووضع هذا العمل أيضًا توقعاتنا بشأن طريقة عملنا مع الموردين، من التطوير إلى الاختبار وإطلاق خدمة تشغيلية [43].

نموذج بيانات الخدمة:

لكي تعمل الخدمة بفاعلية بوصفها كلاً متكاملًا، فإنها تحتاج إلى نموذج بيانات أساسي، لا لإتاحة ملء الباحثين للنماذج لاستيعاب البيانات وتقديم الحد الأدنى من البيانات الوصفية اللازمة لمعرّف (DOI) من (DataCite) فحسب، بل أيضًا لتمكين النظم من إنشاء الأحداث والرسائل ودفعها إلى بعضها بعضًا لتحقيق هدف النظام المتكامل. ويتيح نموذج البيانات أيضًا إثراء البيانات الوصفية وتوليدها آليًا بإتاحة الربط بالنظم المؤسسية والنظم الخارجية للاتصال العلمي.

وقد أُعدّ نموذج بيانات للنسخة ألفا من الخدمة لطرحة للتشاور. [44] وبُنيت هذه المسودة الحالية من نموذج البيانات وفق مطلب التشغيل البيئي مع أكبر عدد ممكن من المنتجات والخدمات التي حُدّدت في حزم الموردين أو في غيرها. فعلى سبيل المثال، نموذج البيانات متوائم مع متطلبات البيانات الوصفية لخدمات الاتصال العلمي.

ونتيجة هذا العمل نموذج بيانات كثيف ومعقد، يضم كيانات وعلاقات ومفردات كثيرة. ودور التشاور هو ضبط هذا النموذج ضبطًا دقيقًا من منظور عملي، وضمان قابليته للتشغيل البيئي مع البنية التحتية القائمة. ويعني ذلك إمكان حذف الخصائص إن لم تُستخدم في النظام الموصوف. ثم يمكن تحديد نموذج بيانات مختصر (ومفردات مختصرة) بوصفه النواة، مع توسيع البنية حسب الحاجة.

التقدم والتحديات:

تخوض (Jisc) الآن مرحلة التطوير ألفا، وتُرسى أسس الخدمة. ويشمل ذلك إعداد البنية التقنية، وإشراك جميع موردي نظم المستودعات والحفظ في بيئة اختبار. وتضم بيئة الاختبار هذه أيضًا نظم بحث واسعة الاستخدام، مثل (EPrints) و (DSpace)، إلى جانب بيانات اختبارية، كي يتمكن موردونا ومطورونا من العمل معًا لتحقيق التشغيل البيئي مبكرًا لمؤسساتنا التجريبية. وعُيّن قائد لتجربة المستخدم لوضع مبادئ وإطار للحوكمة على امتداد المشروع. وسيستمر الاستكشاف وجمع المتطلبات والاختبار مع المؤسسات التجريبية، إلى جانب مدخلات المجتمع من القطاع عبر ورش العمل. [45]

ومن الأعمال الأخرى الجارية حاليًا ما يتعلق بتقارير التكاليف ونمذجة الأعمال، كي تتمكن (Jisc) من تقديم عرض جذاب ماليًا لقطاع التعليم العالي للخدمة الإنتاجية. وإلى جانب ذلك، تُجرى أيضًا دراسة للسوق لاستكشاف احتياجات المؤسسات من خارج المجموعة التجريبية، ومعرفة ما إذا كان نظام لإدارة بيانات البحث تستضيفه (Jisc) جذابًا لها.

ومن المقرر أن ينتهي التطوير ألفا في صيف 2017، وسيتناول التطوير بيتا قابلية النظام للتوسع، ومزيدًا من حالات التكامل، وقضايا صعبة مثل إدارة مجموعات البيانات الضخمة، وإدارة مجموعات البيانات الحساسة، والتحديات المتعلقة بحفظ مجموعة شديدة التنوع من الكائنات الرقمية التي تقع خارج اختصاص مجتمع الحفظ الرقمي التقليدي. ومن المقرر أن ينتهي التطوير بيتا في أبريل 2018؛ غير أن دراسة جدوى لخدمة تشغيلية ستُعرض على مجلس إدارة (Jisc) بحلول ديسمبر 2017.

وطوال مرحلتَي ألفا وبيتا، تلقى مشروع الخدمة متطلبات صعبة من مؤسساتنا التجريبية، منها:

- تحديد «الحد الأدنى من المنتج القابل للاستخدام» (minimum viable product) في ظل تعدد النظم والأولويات والتوقعات
 - الملاءمة مع مسارات العمل القائمة لدى المؤسسات والباحثين، ومن ذلك مثلًا إدراج الخدمة في سياسة مؤسسية تجعل نظام (CRIS) البوابة الأمامية للباحثين
 - جعل الحفظ ملائمًا لبيانات البحث، في حين أن تطوير النظم والأدوات قاده قطاع التراث الثقافي
 - إدارة البيانات الضخمة، أي البيانات الأكبر من أن تُرفع عبر الويب، ومن ثمّ التي يزيد حجمها على 5 غيغابايت، بما في ذلك تحديات البيانات الضخمة (big data)
 - إدارة البيانات الحساسة، ومنها البيانات التجارية، والمعلومات التي تحدد هوية الأشخاص، والبيانات الطبية.
- ويمكن الاطلاع على مزيد من التفاصيل عن المشروع في التحديثات المنتظمة على المدونة [46]، والمواد المرفوعة على موقع شبكة بيانات البحث (Research Data Network). [47]

المراجع:

1. EPSRC Scope and Benefits: <https://www.epsrc.ac.uk/about/standards/researchdata/scope/> (accessed 26 January 2017).

2. Brown, M L and White, W (2013). A partnership approach to research data management In: Pryor, G, Jones, S and White, A eds. Delivering Research Data Management Services: Fundamentals of Good Practice. London: Facet. <http://eprints.soton.ac.uk/id/eprint/356247> (accessed 26 January 2017).
3. Jisc research data shared service: <https://www.jisc.ac.uk/rd/projects/research-data-shared-service> (accessed 26 January 2017).
4. Jisc Research at Risk: <https://www.jisc.ac.uk/rd/projects/research-at-risk> (accessed 22 February 2017).
5. EPSRC policy framework on research data: <https://www.epsrc.ac.uk/about/standards/researchdata/> (accessed 26 January 2017).
6. Research Excellence Framework 2014: <http://www.ref.ac.uk/> (accessed 26 January 2017).
7. Research Councils UK (دون تاريخ). Concordat on Open Research Data launched: <http://www.rcuk.ac.uk/media/news/160728/> (accessed 26 January 2017).
8. Universities UK: <http://www.universitiesuk.ac.uk/> (accessed 26 January 2017).
9. Jisc Managing Research Data programme: <https://www.jisc.ac.uk/rd/projects/managing-research-data> (accessed 26 January 2017).
10. Jisc research data spring: <https://www.jisc.ac.uk/rd/projects/research-data-spring> (accessed 26 January 2017).
11. University of Leeds Research Data Management Pilot Roadmap (). Project Outputs: <https://library.leeds.ac.uk/roadmap-project-outputs> (accessed 26 January 2017).
12. Parsons, T and Berry, M (2012). Research Data Management Technical Requirements: A report to ADMiRe and IS stakeholders, Nottingham, The University of Nottingham: <https://admire.jiscinvolve.org/wp/files/2013/05/ADMiRe-RDM-Technical-Requirements-Report.pdf> (accessed 26 January 2017).
13. Garret, L, Silva, C and Gramstadt, M-T (2011). Kaptur technical analysis report, University for the Creative Arts: https://vads.ac.uk/kaptur/outputs/Kaptur_technical_analysis.pdf (accessed 26 January 2017).
14. Jones, R (2012). Sword Data Deposit Scenarios July 3 2012 SWORD: <http://swordapp.org/2012/07/data-deposit-scenarios/> (accessed 26 January 2017).
15. Filling the Digital Preservation Gap: <https://www.york.ac.uk/borthwick/projects/archivematica/> (accessed 26 January 2017).

16. Miller, A (2015). Project Report: A consortial approach to integrated RDMS In: figshare. <https://dx.doi.org/10.6084/m9.figshare.1480451.v1> (accessed 26 January 2017).
17. Data Vault: <http://libraryblogs.is.ed.ac.uk/jiscdatavault/> (accessed 26 January 2017).
18. Effective learning analytics: <https://www.jisc.ac.uk/rd/projects/effective-learning-analytics> (accessed 26 January 2017).
19. Kaye, J (). Jisc RDM Shared Service Pilot Initial Statement of Requirements: <https://researchdata.jiscinvolve.org/wp/files/2015/11/Draft-RDMSS-Requirements-Specification-V1.0.docx> (accessed 26 January 2017).
20. Duca, D (2015). What makes up the 'ideal' research data management system? July 9 2015 Jisc Shared Services Workshop: <https://researchdata.jiscinvolve.org/wp/2015/07/30/makes-ideal-research-data-management-system/> (accessed 26 January 2017).
21. Lewis, J A (2014). Research Data Management Technical Infrastructure: A Review of Options for Development at the University of Sheffield In: figshare. <https://dx.doi.org/10.6084/m9.figshare.1202230.v9> (accessed 26 January 2017).
22. RDM Shared Services November Workshops: <https://researchdata.jiscinvolve.org/wp/2015/11/23/rdm-shared-service-workshops/> (accessed 26 January 2017).
23. Kaye, J, Stokes, P and Bruce, R (). Jisc Research Data Shared Service Operational Requirements Zenodo: <http://doi.org/10.5281/zenodo.48261> (accessed 26 January 2017).
24. Research Data Management Shared Service - Call for Formal Expressions of Interest <https://researchdata.jiscinvolve.org/wp/2015/11/06/research-data-management-shared-service-call-for-formal-expressions-of-interest/> (accessed 26 January 2017).
25. Kaye, J, Stokes, P and Bruce, R (). Jisc Research Data Shared Service Operational Requirements Zenodo: <http://doi.org/10.5281/zenodo.48261> (accessed 26 January 2017).
26. Kaye, J (). Research Data Shared Service - OR2016: <https://researchdata.jiscinvolve.org/wp/2016/06/14/jisc-research-data-shared-service-or2016/> (accessed 26 January 2017).
27. Research Consulting: <http://www.research-consulting.com/> (accessed 26 January 2017).
28. Data Asset Framework (DAF): <http://www.data-audit.eu/index.html> (accessed 26 January 2017).
29. ORCID: <https://orcid.org/> (accessed 14th January 2017).
30. DataCite: <https://www.datacite.org/> (accessed 26 January 2017).
31. DAF (). ref. 28.

32. Research Consulting: 2016 DAF Survey Results: <https://researchdata.jiscinvolve.org/wp/files/2016/11/2016-DAF-survey-results-for-blog.pptx> (accessed 26 January 2017).
33. Johnson, R, Parsons, T, Chiarelli, A and Kaye, J (2016). Jisc Research Data Assessment Support - Findings of the 2016 data assessment framework (DAF) surveys DOI: <https://doi.org/10.5281/zenodo.177856> Zenodo; (accessed 26 January 2017).
34. Johnson, R, Chiarelli, A and Parsons, T (2016). Data asset framework (DAF) survey results In: figshare. <http://dx.doi.org/10.6084/m9.figshare.3796305.v4> (accessed 26 January 2017).
35. Clax Ltd. <http://www.clax.co.uk/> (accessed 26 January 2017).
36. Ferguson, N (2016). Report for the proposed Research Data Shared Service on focus groups held between May and October 2016 and the metadata issues and requirements identified DOI: <https://doi.org/10.5281/zenodo.193018> Zenodo; (accessed 14 January 2017).
37. Ferguson, N (2016). Jisc Research Data Shared Service metadata focus group use cases [data set] DOI: <https://doi.org/10.5281/zenodo.193011> Zenodo; (accessed 26 January 2017).
38. Kaye, J and Bruce, R (). Research data shared service, Poster to Open Repositories 2016: https://researchdata.jiscinvolve.org/wp/files/2016/06/RDSS_POSTER_JUNE2016_FINAL.pdf (accessed 26 January 2017).
39. UK Research Data Discovery Service (Alpha) (). <http://ckan.data.alpha.jisc.ac.uk/dataset> (accessed 26 January 2017).
40. IRUS for Data [data set]: <https://www.jisc.ac.uk/rd/projects/research-data-metrics-for-usage> (accessed 26 January 2017).
41. UK ORCID consortium (). <https://www.jisc.ac.uk/orcid> (accessed 26 January 2017).
42. Meeting the requirements of the EPSRC research data policy: <https://www.jisc.ac.uk/guides/meeting-the-requirements-of-the-EPSRC-research-data-policy> (accessed 26 January 2017).
43. RDSS Conceptual Technical Architecture: <https://www.lucidchart.com/documents/edit/6398e8e5-51fb-46ff-8e08-7f97e8861265> (accessed 16 January 2017).
44. RDSS Canonical Data Model: <https://github.com/JiscRDSS/rdss-canonical-data-model> (accessed 26 January 2017).
45. Research Data Network: <http://researchdata.network> (accessed 26 January 2017).

46. Jisc RDM Blog: <https://researchdata.jiscinvolve.org/wp/> (accessed 26 January 2017).

47. Research Data Network, ref. 46.