الحفظ الرقمي في بيئة متغيرة: تحديات وفرص من تجربة بورتيكو

كيت ويتنبرغ، سارة غلاسر، إيمي كيرشهوف، شيلا موريسي، ستيفاني أورفان

الملخص:

يتناول هذا المقال، انطلاقًا من تجربة خدمة الحفظ الرقمي بورتيكو، تحديات الحفاظ على المحتوى الرقمي للاتصال العلمي وفرصه. فيستعرض نشأة الخدمة عام 2005 في سياق مخاوف المكتبات من فقدان السيطرة المادية على المحتوى الإلكتروني، وتطور نموذجها من حفظ الدوريات الإلكترونية إلى الكتب الإلكترونية والمجموعات المرقمنة، ونمو المشاركة فيها، وآلية «الأحداث المحفِّزة» لإتاحة المحتوى، والتعاون مع المكتبة البريطانية في الإيداع القانوني الرقمي. ثم يحلل التحديات الراهنة: تراجع التمويل، ونماذج النشر وإعادة الاستخدام الجديدة ومنها الوصول المفتوح وبيانات البحث، وتعقّد الكائنات الرقمية الموزعة. ويعرض الفرص المتاحة في التعاون الواسع والتقدم التقني، كالتخزين السحابي والبرمجيات المفتوحة والمحاكاة وأدوات التنقيب في النصوص والبيانات. ويخلص إلى أن مواكبة التغير تتطلب مرونة وفهمًا عميقًا لمنظومة الاتصال العلمي المتطورة.

الكلمات المفتاحية: الحفظ الرقمي، بورتيكو، الدوريات والكتب الإلكترونية، الإيداع القانوني الرقمي، التعاون بين المكتبات والناشرين

مقدمة:

شهدت العقود الثلاثة الأخيرة نموًا هائلًا في كمية المحتوى الرقمي الذي تنتجه المكتبات والناشرون والمؤسسات الثقافية وعامة الناس. وتنطوي إتاحة المحتوى في شكل رقمي على فوائد كبيرة. غير أن الكائنات الرقمية، على خلاف الكائنات المطبوعة التي يمكن أن تدوم عقودًا كثيرة بأدنى قدر من العناية متى طُبعت على ورق خالٍ من الأحماض وحُفظت في ظروف معقولة، قد تكون قصيرة العمر للغاية ما لم يُولَ حفظها العناية اللازمة. ويُعدّ الحفظ طويل الأمد لهذا المحتوى الرقمي شاغلًا رئيسيًا للمجتمع العلمي ونحن نواصل الانتقال من المطبوع إلى الرقمي.

وتقدّم مؤسسات كثيرة خدمات الحفظ الرقمي، منها الهيئات الحكومية، والمنظمات غير الربحية التي تقدّم خدمات للحفظ، والمكتبات، والجهات التجارية المهتمة بحفظ محتواها لأغراض البحث في المستقبل. وقد مضى أكثر من 15 عامًا على إنشاء تحالف الحفظ الرقمي (Digital Preservation Coalition) وعلى المرحلة المبكرة لخدمات يدعمها المجتمع مثل بورتيكو و(LOCKSS)، وتعلّمنا خلال هذه المدة الكثير عن التحديات القائمة في توفير حفظ طويل الأمد موثوق وقابل للتوسع ومستدام، وفي توفير الوصول إلى المحتوى المحفوظ عند الحاجة.

وانطلاقًا من هذه التجربة، ومن الحاجة المستمرة إلى الموارد والاهتمام بالحفظ الرقمي، نقدّم هنا خلفية عن تاريخ بورتيكو وعن ممارستها المتطورة في الحفظ المستدام للمنتجات الرقمية للاتصال العلمي. ونقدّم كذلك عرضًا عامًا لمشهد الحفظ الرقمي كما نراه الآن، مع أفكارنا حول المتطلبات الحالية والفرص والتحديات التي تنتظرنا.

نظرة إلى الوراء: بورتيكو بوصفها دراسة حالة:

في عام 2003 كانت المكتبات والناشرون على السواء يعانون تداعيات إفلاس (RoweCom/Faxon). ولم يكن قد مضى عام بعد على صدور تقرير (RLG/OCLC) بعنوان «Trusted Digital Repositories: Attributes and Responsibilities» (المستودعات الرقمية الموثوقة: السمات والمسؤوليات)[1]. وبحلول عام 2005، وهو العام الذي شهد الإطلاق الرسمي لبورتيكو، صدر تقرير مؤسسة أندرو و. ميلون (Andrew W Mellon Foundation) بعنوان «Urgent Action Needed to Preserve Scholarly Electronic Journals» (الحاجة إلى إجراء عاجل لحفظ الدوريات العلمية الإلكترونية)[2]، وحظي بتأييد واسع من مجتمع المكتبات الأكاديمية. وفي هذا السياق ظهرت خدمات الحفظ التي يقدّمها طرف ثالث مثل بورتيكو، استجابةً للاحتياجات التي عبّر عنها مجتمع المكتبات الأكاديمية، إذ إن تزايد اعتماده على الدوريات الإلكترونية وغيرها من الموارد الإلكترونية قد يسّر الوصول إلى المحتوى، لكنه اقتضى التخلي عن السيطرة المادية التقليدية على صونه.

تطور نموذج الخدمة: الكتب الإلكترونية والدوريات الإلكترونية والمجموعات المرقمنة:

اقتصر نطاق عمل بورتيكو في سنواتها الأولى على حفظ محتوى الدوريات الإلكترونية، وكان الهدف النهائي تمكين المجتمع الأكاديمي من الوصول إلى ذلك المحتوى إذا لم يعد متاحًا على الإنترنت لدى الناشر أو من يخلفه (وهو ما يُسمّى في مصطلحات بورتيكو «الحدث المحفِّز» (trigger event)). وتقاسم مجتمعا المكتبات الأكاديمية والنشر العلمي تكلفة القيام بهذا العمل المهم، أي الحفظ، فضمنا بذلك حفظ طيف واسع من المحتوى، واستدامة الخدمة على المدى الطويل. وبحلول عام 2009 بدأت بورتيكو حفظ محتوى الكتب الإلكترونية، وفق نموذج مجتمعي أيضًا، وأطلقت كذلك خدمة خاصة بالمجموعات التاريخية المرقمنة، يدعمها الناشرون وحدهم.

تطور الدعم المجتمعي:

في عام 2006 كانت بورتيكو تحظى بدعم 27 ناشرًا و245 مكتبة. ويشارك في بورتيكو حاليًا 518 ناشرًا (يمثلون أكثر من 2,000 جمعية علمية ورابطة). ويضم المحتوى الذي التزم به هؤلاء الناشرون 28,468 عنوانًا من الدوريات الإلكترونية، و1,231,039 عنوانًا من الكتب الإلكترونية، و187 مجموعة مرقمنة. ونما دعم المكتبات كذلك نموًا مطردًا خلال العقد الماضي. واليوم يشارك في بورتيكو أكثر من 1,000 مكتبة في أنحاء العالم.

وما زالت الدوافع الأصلية لخدمات الحفظ التي يقدّمها طرف ثالث قائمة. بل إنها، إن كان ثمة تغيير، قد اشتدت مع الزيادة الهائلة في اتساع المحتوى العلمي المتاح على الإنترنت وعمقه. فالمكتبات تختار على نحو متزايد الوصول الإلكتروني وحده إلى الدوريات والكتب، متخليةً عن المطبوع كليًا. وكثير من الموارد لم يعد يتيحه مقدّمو المحتوى إلا في شكل إلكتروني. والتحول من المحتوى الرقمي بوصفه طريقة وصول ثانوية إلى المطبوع، إلى المحتوى الرقمي بوصفه النسخة المعتمدة (version of record) وطريقة الوصول الأساسية، يجعل الحاجة أشدّ إلى حفظ متين يقدّمه طرف ثالث، وإلى ما يوفّره من ضمان للوصول في المستقبل.

وفي نهاية المطاف، توجد خدمات الحفظ لضمان أن يظل المحتوى العلمي المتاح على الإنترنت اليوم متاحًا دائمًا في المستقبل. ويقوم دعم المكتبات لخدمات الطرف الثالث ومشاركتها فيها على الوعد بإمكانية الاستخدام المستمرة والوصول في المستقبل عند الضرورة. وتفي بورتيكو وغيرها من خدمات الحفظ بهذا الوعد من خلال الأحداث المحفِّزة، وهي الآلية التي يُتاح بها المحتوى حين يصبح، كما ذُكر، غير متاح لدى الناشر الأصلي أو من يخلفه.

وتقع الأحداث المحفِّزة لأسباب عدة، تتراوح بين خروج ناشر من السوق، واستحواذ ناشر آخر عليه دون نقل عناوينه المتوقفة إلى منصة الناشر المستحوِذ، وقرار ناشر صراحةً التوقف عن استضافة منشور ما لنقص الدعم المالي. وفي نموذج الخدمة الأصلي لبورتيكو، كان المحتوى الذي فُعّلت إتاحته (triggered content) يُتاح للمكتبات المشاركة في بورتيكو وحدها (بصرف النظر عما إذا كانت قد اشتركت يومًا في عنوان المنشور الذي فُعّلت إتاحته). ومع مرور الوقت، ومع انتقال الوصول المفتوح (open access, OA) من نموذج نشر هامشي إلى نموذج شائع، غيّرت بورتيكو سياستها تلبيةً لتوقع المجتمع أن يظل المحتوى المفتوح مفتوحًا، وصرنا نمكّن ناشري الوصول المفتوح من ضمان بقاء محتواهم مفتوحًا عند تفعيل إتاحته.

وحتى الآن، فعّلت بورتيكو إتاحة محتوى 46 دورية إلكترونية، 25 منها من دوريات الوصول المفتوح. وأتحنا كذلك عملًا مرجعيًا وقاعدة بيانات للمستخلصات. وإلى جانب إتاحة المحتوى نتيجةً للأحداث المحفِّزة، تتيح بورتيكو أيضًا، باتفاق مع بعض ناشريها، المحتوى بصفتها مزوّدًا للوصول بعد إلغاء الاشتراك/الوصول الدائم (post-cancellation/ perpetual access) لعدد كبير من العناوين التي تحفظها. وحتى ديسمبر 2017، أُتيح محتوى 424 دورية و404 كتب لـ251 مؤسسة من خلال الوصول بعد إلغاء الاشتراك/الوصول الدائم.

ومع أن محتوى أرشيفنا كله قد نما، تظل الدوريات الإلكترونية خدمتنا الأساسية، ومع أن دوريات كثيرة محفوظة في بورتيكو، فإن عدد عناوين الدوريات القائمة يفوق بكثير ما تحفظه الخدمة. ويأتي عدد كبير منها من ناشرين أصغر (يُعرَّفون بأنهم من ينشرون عشرة عناوين أو أقل). ويمكن القول إن هذه الدوريات هي الأكثر عرضة للاختفاء إن لم تُحفظ. غير أن إدخال هذه الدوريات في خدمات مثل بورتيكو يمثل تحديًا أكبر. فأولًا، على الرغم من انخفاض الرسوم المفروضة على هؤلاء الناشرين الأصغر (إذ تُحسب الرسوم على أساس الإيرادات)، قد تكون حتى الرسوم الرمزية عائقًا ماليًا أمام ناشر محدود الإيرادات وقليل الدعم المؤسسي. وثانيًا، بعضهم صغير إلى حدّ أن القيود التنظيمية وقيود الموارد تجعل مشاركتهم صعبة. فعلى سبيل المثال، يفتقر كثير منهم إلى إجراءات راسخة (تقنية وإدارية على السواء) لاعتماد إيداعات المحتوى لغرض الحفظ وإدارتها. وقد تمكنت بورتيكو من إحراز تقدّم في هذا المجال بتطوير أدوات لدعم الناشرين الأصغر (مثل إضافة برمجية لتصدير المحتوى لنظام (Open Journal Systems))، وبتطوير آليات أكثر مرونة للحصول على البيانات الوصفية وإيداعات المحتوى. وتحفظ بورتيكو حاليًا محتوى أكثر من

200 ناشر صغير. ومع أن ذلك يمثل تقدمًا جيدًا، ما زال أمامنا عمل كثير في هذا المجال؛ وعلى سبيل الإشارة، نعلم أن ثمة ما لا يقل عن 7,000 ناشر مدرجين في (Crossref) ليس لهم محتوى محفوظ في بورتيكو أو في غيرها من خدمات الحفظ، ونقدّر أن 4,700 منهم ناشرون صغار.

تطور نموذج الخدمة: التعاون مع المكتبات الوطنية:

من أجل الاستفادة القصوى من جهود حفظ المواد العلمية، ثبت أن من المفيد للمؤسسات أن تستكشف سبل تنسيق جهودها والتعاون فيما بينها. ومن أمثلة هذا التعاون العلاقة بين بورتيكو والمكتبة البريطانية (British Library).

ففي أبريل 2013 طبّقت المملكة المتحدة تشريعًا يوجب الإيداع القانوني للمحتوى الرقمي المنشور فيها، وتضطلع المكتبة البريطانية بدور مكتبة الإيداع القانوني ومنسّق البرنامج. والتشريع واسع النطاق، إذ يشمل «الكتب الإلكترونية والدوريات الإلكترونية وغيرها من أنواع المنشورات الإلكترونية، إضافةً إلى المواد الأخرى التي تُتاح للجمهور في المملكة المتحدة على وسائط محمولة مثل الأقراص المدمجة (CD-ROM) والمصغّرات الفيلمية، وعلى الويب (بما في ذلك المواقع الإلكترونية)، وبالتنزيل من موقع إلكتروني» [3]. ويغطي أحد أجزاء تشريع الإيداع القانوني في المملكة المتحدة الدوريات الإلكترونية، وكان كثير منها يُودَع ويُحفظ أصلًا في بورتيكو. وفي عام 2013 أبرمت بورتيكو اتفاقًا مع المكتبة البريطانية لتقديم خدمة لمعالجة الدوريات الإلكترونية تدعم متطلبات الإيداع القانوني لدى المكتبة. وتتيح هذه العلاقة للمكتبة البريطانية الاستفادة من البنية التحتية التقنية القائمة لبورتيكو ومن خبرة العاملين فيها. فبورتيكو تتعامل أصلًا مع تدفقات محتوى متنوعة من الناشرين، فتوحّد معايير البيانات الوصفية للمقالات، وتعيد تنظيم المقالات والأعداد في حزمة تلائم نموذج محتوى الحفظ لدينا. ثم يُصدَّر هذا المحتوى «المُوحَّد» (normalized) من أنظمة بورتيكو إلى المكتبة البريطانية، فيوفر لها تدفقًا واحدًا من المحتوى تديره.

وحتى مايو 2018، تزوّد بورتيكو المكتبة البريطانية بمحتوى الدوريات الإلكترونية الجارية من 21 ناشرًا، يضم قرابة 7,000 دورية، وقد سُلّم منه أكثر من 3.7 ملايين مقالة. وتستوعب المكتبة البريطانية المحتوى الوارد من بورتيكو في نظام الحفظ الخاص بها، ثم ترسله إلى مكتبات الإيداع في أنحاء المملكة المتحدة وأيرلندا، حيث تُقدَّم مقالات الدوريات الإلكترونية إلى روّاد المكتبات في قاعات المطالعة.

ومع تزايد حجم المحتوى الذي يحتاج إلى الحفظ وتعقيده، ستكون هذه الأنواع من العلاقات التعاونية أساسية لتلبية احتياجات المجتمع من الحفظ بتكلفة فعالة.

نظرة حولنا: مشهد الاتصال العلمي:

بالنظر إلى الاحتياجات والتحديات التي دفعت إلى إنشاء بورتيكو قبل 15 عامًا، ما بعض التحديات والفرص القائمة في مشهد الحفظ الرقمي الحالي والمستقبلي؟

التحدي: الاقتصاد:

من التحديات الدائمة التحدي الاقتصادي، ولا سيما لمؤسسات التعليم العالي المعنية بحفظ الدوريات العلمية والدراسات المفردة (monographs) والمجموعات الخاصة. فثمة اتجاه مستمر نحو انخفاض الأموال العامة المخصصة لدعم التعليم العالي، يصاحبه تزايد التسييس في النقاشات حول سياسات التعليم وممارساته. وتطال الآثار المتتالية لهذا الاتجاه المكتباتِ البحثية التي اضطرت إلى العمل بموارد متناقصة، فصارت قدرتها على دعم خدمات الحفظ مقيدة على نحو متزايد. فعلى سبيل المثال، لا تشارك غالبية المكتبات الجامعية في أي خدمة للحفظ الرقمي. وقد أدى الضغط لتقليص ميزانيات المكتبات إلى اضطرار كثير من المؤسسات إلى توجيه أموالها نحو اقتناء المحتوى بدلًا من حفظه، في حين قد ترى المكتبات الأصغر أن الحفظ خارج عن نطاق تكليفها المؤسسي.

التحدي: نماذج جديدة للنشر وإعادة الاستخدام:

أدت المكتبات الأكاديمية والبحثية نفسها على نحو متزايد دور الناشر، بنشر «الدوريات الإلكترونية، ووقائع المؤتمرات، والتقارير الفنية، والمواقع الإلكترونية القائمة على قواعد البيانات» [4]. والأشدّ تحديًا، من منظور الحفظ، دور المكتبات بوصفها قيّمة على ما يمكن تسميته «المحتوى المُنشأ محليًا»: المستودعات المؤسسية (institutional repositories) وغيرها من الأنظمة التي تضم مجموعات متخصصة، ومواد تعليمية، ونسخًا أولية (preprints)، وكائنات أخرى. ومع أن المكتبات البحثية والوطنية الأكبر تخصص عمومًا في ميزانياتها اعتمادات لخدمات الحفظ (للمحتوى الداخلي والخارجي على السواء)، فإن كثيرًا من المؤسسات الأصغر تملك مجموعات خاصة نادرة وفريدة تحتاج إلى الحفظ. ومن دون تمويل كافٍ لهذه المكتبات والمؤسسات الثقافية الأصغر، لن تُحفظ هذه المجموعات الخاصة، مما سيعرّض هذا المحتوى القيّم للخطر على المدى الطويل.

ويجد الناشرون التقليديون أن عليهم الاستجابة لضغط متزايد من أجل الوصول المفتوح إلى منتجات البحث العلمي. وقد ازداد الدفع نحو الوصول المفتوح زيادة ملحوظة بفضل مبادرة بودابست للوصول المفتوح (Budapest Open Access initiative) عام 2001 [5]، وبيان بيثيسدا (Bethesda Statement) عام 2003 [6]، وإعلان برلين بشأن الوصول المفتوح إلى المعرفة في العلوم والإنسانيات (Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities) [7]، ونشر تقرير فينش (Finch Report) في يونيو 2012 [8]، ومذكرة مكتب سياسات العلوم والتكنولوجيا (Office of Science and Technology Policy, OSTP) الصادرة في فبراير 2013 بعنوان «(Expanding Public Access to the Results of Federally Funded Research)» (توسيع وصول الجمهور إلى نتائج البحوث الممولة اتحاديًا)[9]. غير أنه تجدر الإشارة إلى أن أولويات الحفظ تميل إلى التمحور حول المحتوى الذي تدفع المكتبات مقابل الوصول إليه، مع أن محتوى الوصول المفتوح يحتاج هو أيضًا إلى الحفظ.

ونتيجةً جزئيةً لمبادرات توسيع وصول الجمهور إلى نتائج البحث، مثل مذكرة (OSTP) ومبادئ (FAIR) التوجيهية لإدارة البيانات العلمية ورعايتها (FAIR Guiding Principles for scientific data management and stewardship) [10]، ثمة أيضًا تركيز متزايد على جمع مجموعات البيانات وغيرها من منتجات البحث وحفظها ومشاركتها، وعلى إعداد خطط لإدارة البيانات ضمن جميع مقترحات المنح تحقيقًا لذلك. والحفظ، الذي تعرّفه بورتيكو بأنه توفير إمكانية استخدام المحتوى وأصالته وقابليته للاكتشاف وإمكانية الوصول إليه على المدى الطويل، لا بدّ له من أن يتعامل مع قضايا إتاحة البيانات لإعادة الاستخدام، ومنها جعل مجموعات البيانات الكبيرة أو

المعقدة مفهومة عبر التصورات المرئية وغيرها من أشكال التمثيل، ومع قضايا ضمان قابلية اكتشاف المنتجات غير النصية مثل الصور والفيديو والنماذج الهندسية والمعمارية ونظم المعلومات الجغرافية وغيرها من صيغ البيانات المتخصصة.

التحدي: تعقّد الكائنات الرقمية:

والواقع أنه إلى جانب مجموعات البيانات والمنتجات غير النصية، أصبح التعقّد المتزايد للمحتوى الرقمي تحديًا قائمًا بذاته. فكما يصف تقرير (OCLC) الصادر عام 2014 بعنوان «(The Evolving Scholarly Record)» (السجل العلمي المتطور)[11]، أصبح المحتوى الذي يتألف منه السجل العلمي، مقارنةً بأشكاله السابقة المكتوبة والمطبوعة بل حتى الرقمية المبكرة، أكثر تغيرًا وأقل وضوحًا من حيث الحدود. ففي السابق، كان حتى المنتج الرقمي من السجل العلمي يتألف من كائن منفصل إلى حدّ ما، مثل مقالة في دورية أو كتاب، وكثيرًا ما كان محتوًى في ملف واحد. أما اليوم فيرجّح على نحو متزايد أن يكون هذا المنتج ما صرنا نسميه «كائنًا علميًا مركبًا موزعًا» (distributed scholarly compound object). فقد توجد مكوناته المختلفة في أكثر من مستودع واحد (بما في ذلك مستودعات النسخ الأولية)، وقد تكون لها أكثر من نسخة، تعتمد على طبقات معقدة من منصات الأجهزة والبرمجيات المتغيرة باستمرار. ولكي تشكّل هذه الكائنات سجلًا علميًا، ولكي يُحفظ ذلك السجل، يتعيّن على أرشيفات الحفظ أن تنشئ وتجمع وتصون معلومات عن السياق والعلاقات أكثر حتى مما توفره البيانات الوصفية البيليوغرافية «التقليدية». ويفضي هذا المحتوى أيضًا إلى متطلبات متزايدة التعقيد في الاكتشاف (discovery) والوصول والعرض.

ومما يزيد تحديات الرعاية (stewardship) تعقيدًا ما يسميه كليفورد لينش (Clifford Lynch) «الأنظمة المدفوعة بالخوارزميات» (algorithmically driven systems) [12]، مثل تطبيقات وسائل التواصل الاجتماعي الشائعة، التي باتت تشكّل هي أيضًا عُقدًا في الشبكة المعقدة الجديدة للاتصال العلمي.

كيف نواجه هذه التحديات:

الفرصة: التعاون:

من الاستجابات الرئيسية لمجتمع الحفظ الرقمي التعاون، على نطاق واسع وعبر الحدود التخصصية والمهنية والوطنية. فبورتيكو نفسها قائمة على تعاون من المكتبات والناشرين معًا لضمان الحفظ طويل الأمد للدوريات والكتب العلمية المرقمنة والمنشأة رقميًا، والمواد المرجعية، ومجموعات المصادر الأولية. وتعتمد بورتيكو على مجتمع الممارسة الدولي، وتسهم فيه في آن واحد، وهو مجتمع منخرط باستمرار في التجريب، وتحديد المعايير وأفضل الممارسات، وتطوير مجموعات الأدوات والأطر اللازمة للحفظ. وعملنا مع المكتبة البريطانية مثال على كيفية جمعنا بين عمل مؤسستين مستقلتين متقاربتين في التوجه لخدمة شراكة تعاونية تعود بالنفع على البحث العلمي.

وعلى نحو أعم، اتخذت المكتبات الجامعية والبحثية إجراءات جماعية لمواجهة تحديات الحفظ الرقمي. ففي الولايات المتحدة، تشمل مبادرات الحفظ التعاونية البارزة شبكة الحفظ الرقمي (Digital Preservation Network)، و(CLOCKSS)، وشبكات (LOCKSS) الخاصة (Private LOCKSS Networks) المختلفة. كما تتلقى بورتيكو على نحو متزايد طلبات من اتحادات المكتبات (مثل الاتحاد السويسري (Swiss Consortium) ، ومؤسسة البحث الألمانية (Deutsche Forschungsgemeinschaft) في ألمانيا، و(CAPES) في البرازيل) التي تبحث عن استراتيجيات وطنية

للحفظ الرقمي. وتعاونت بورتيكو إضافةً إلى ذلك مع شركاء في مشروعات بحثية في مجال الحفظ، منها مبادرة حديثة مدتها عامان مولتها مؤسسة ألفريد ب. سلون (Alfred P Sloan Foundation)، نُفذت مع (Data Conservancy) التابعة لمكتبة جامعة جونز هوبكنز (Johns Hopkins University Library Data Conservancy) ومع (IEEE)، لتطوير وسائل حفظ العلاقات المعقدة بين المنشورات العلمية والبيانات التي تستند إليها، دعمًا للتطور المستمر للاتصال العلمي والنشر الرقمي[13].

الفرصة: التقدم التقني:

اتسع مجتمع الممارسة الدولي للحفظ كذلك إلى ما وراء مركزه في المكتبات والأرشيفات، فنتجت عن ذلك منظورات وممارسات وأدوات وتقنيات جديدة. فحاجة المجتمع العلمي، على سبيل المثال، إلى ضمان قابلية إعادة إنتاج النتائج التجريبية، تستند إلى مناهج مؤسسات الحفظ «التقليدية» وتُثريها في آن واحد. ومشاركة المؤسسات التجارية والمالية والحكومية في تحالف الحفظ الرقمي مؤشر على إدراك متسع للحاجة إلى توفير الحفظ للمنتجات الرقمية. وقد أقامت مؤسسات مثل (VIAA) في بلجيكا، وجامعة إنديانا (Indiana University) في الولايات المتحدة ضمن مبادرتها لرقمنة الوسائط وحفظها (Media Digitization and Preservation Initiative)، شراكات جديدة مبتكرة مع القطاع الخاص لتطوير حلول لتحديات حفظ المنتجات السمعية البصرية الموجودة على وسائط وصيغ قديمة.

وشهدت السنوات العشر الأخيرة أيضًا تطورات تقنية مهمة تتبناها مؤسسات الحفظ (ومنها بورتيكو، والمكتبة البريطانية، ومستودع ستانفورد الرقمي (Stanford Digital Repository)، والأرشيف الوطني البريطاني (UK National Archives) التي شرعت مؤخرًا، أو ستشرع قريبًا، في هندسة الجيل التالي من بنيتها التحتية التقنية للحفظ. فقد اتسعت باقة تقنيات بنية الحفظ لتشمل منصات التخزين السحابي والخوادم التجارية وغير التجارية (مثل (AWS من (Amazon) [14] و(Azure) من (Microsoft) [15]). ولعل الأهم من ذلك أن كثيرًا من المكتبات البرمجية وواجهات برمجة التطبيقات (APIs) التي يستخدمها كبار مزوّدي التطبيقات والمنصات التجارية قد أُصدرت برمجياتٍ حرة ومفتوحة المصدر، فوفّرت أدوات جيدة الدعم لقابلية التوسع الأفقي في التخزين والخوادم على السواء (وهو أمر حاسم لإدارة نمو الأرشيفات حجمًا وتعقيدًا على نحو مستدام).

وفي صميم تحدي الوصول طويل الأمد إلى المنتجات الرقمية يكمن التقادم المستمر والسريع للأجهزة والبرمجيات التي استُخدمت أصلًا في إنشاء تلك المنتجات وعرضها. وتوفر تقنية أخرى، هي المحاكاة (emulation)، أداة إضافية للمساعدة في حل مشكلة التقادم التقني هذه. والمحاكاة هي استخدام برمجيات متخصصة (وأحيانًا أجهزة) لتمكين نظام حاسوبي من محاكاة نظام آخر، أي التصرف مثله، بما في ذلك الأجهزة وأنظمة التشغيل والبرامج القديمة التي تتيح الوصول إلى المنتجات الرقمية القديمة. وقد أوجد التقدم في المحاكاة، إلى جانب مشروعات بحثية في جامعة فرايبورغ (University of Freiburg) [16] وجامعة ييل (Yale University) [17]، قدرات جديدة لتوظيف المحاكاة في خدمة الحفظ.

ولتطوير أدوات معالجة اللغات الطبيعية والتعلم الآلي وغيرها من أدوات التنقيب في النصوص والبيانات (text and data mining) على نطاق واسع، وإتاحتها إتاحة مفتوحة، إمكانات لأتمتة استخلاص البيانات الوصفية وإثرائها، وكلاهما قد يكون صعبًا ومكلفًا على نطاق واسع. ويبقى السؤال مطروحًا عمّا إذا كان استخدام هذه الأدوات سيفضي إلى جمع بيانات وصفية بيليوغرافية «جيدة بما يكفي» بطرق قد تخفف على الأقل بعض عبء حفظ الذيل الطويل من الدوريات العلمية. وبدعم من منحة أعلنت عنها مؤسسة ميلون مؤخرًا[18]، سيستكشف أرشيف

الإنترنت (Internet Archive) استخدام بعض هذه الأدوات في تحليل عمليات الحصاد الواسعة للويب، بهدف تخفيف العبء عن مجتمع المكتبات في تحديد أيٍّ من منشورات الذيل الطويل الكثيرة ينبغي أن تحظى بالأولوية والاهتمام في الحفظ، وترتيبها حسب الأولوية.

نظرة إلى الأمام:

يتواصل نمو تحدي الحفظ الرقمي، ولأن المنتجات التي تُجمع وتُحفظ تعتمد على البرمجيات، فإن توفير وصول ذي معنى إلى هذا المحتوى سيزداد تعقيدًا. وكما ذُكر أعلاه، أتاح صعود المناهج البحثية متعددة التخصصات والتعاونية، التي كثيرًا ما تمتد عبر تخصصات متعددة وتدمج النصوص والبيانات والمواد السمعية البصرية، أشكالًا جديدة من الإنتاج العلمي. فما الذي يمكن فعله لجعل بعض هذه المنتجات «المنشأة رقميًا» حقًا «قابلةً للحفظ منذ إنشائها» أيضًا؟

وقد تكون البنية التحتية القائمة حاليًا غير كافية لدعم الأنشطة العلمية الجديدة على طيف يمتد من التعاون غير الرسمي إلى نسخ مستقرة ومفهرسة من العمل صالحة للنشر والحفظ طويل الأمد. وثمة حاجة إلى أن يفهم المجتمع ما يلزم لدعم هذه الأشكال الجديدة من المحتوى في أثناء انتقالها عبر عملية الاتصال العلمي. كيف ينبغي أن نصمم البنية التحتية اللازمة لدعم هذا العمل؟ وكيف نحفظ الإنتاج الأكاديمي الذي ينبثق من أنماط البحث والنقاش الأقل رسمية التي يمارسها الباحثون خارج دورة النشر التقليدية؟ إن استعداد المجتمع وقدرته على تقديم التوجيه والدعم اللذين يحتاج إليهما الباحثون سيؤديان دورًا مهمًا في تمكين إنشاء هذا العمل ونشره وحفظه في المستقبل.

فما الذي يوحي به ماضي بورتيكو ونحن نتأمل هذا المستقبل؟ مع نمو عدد المكتبات والناشرين المشاركين في خدمات الحفظ، رأت هذه المجتمعات بنفسها أن الاستثمار فيما عُدّ في الأصل تجربةً قد أثمر خدمة حفظ متينة واسعة النطاق ومستدامة اقتصاديًا يقدّمها طرف ثالث. وما تعلمناه في بورتيكو، خلال ما نعدّه تجربة مستمرة، هو أن طبيعة المحتوى الذي نحفظه، والتقنية التي تدعم الخدمة، ونموذج العمل وخصائص الخدمة، وشبكات الشراكات وعلاقاتنا بأولئك الشركاء، قد تطورت كلها، ولا بدّ أن تواصل التطور وستواصله. ومع أن رسالتنا ما زالت كما كانت عند نشأة بورتيكو، فإن القدرة على مواكبة وتيرة التغير تتطلب سرعة الاستجابة والمرونة وفهمًا عميقًا للمنظومة العلمية المتطورة.

المراجع:

1. Trusted Digital Repositories: Attributes and Responsibilities, RLG-OCLC Report, 2002: https://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf (accessed 23 May 2018).
2. Waters D, Urgent Action Needed to Preserve Scholarly Electronic Journals, Andrew W Mellon Foundation, 2005: https://old.diglib.org/pubs/waters051015.pdf (accessed 23 May 2018).
3. The British Library legal deposit for websites and electronic publications: http://www.bl.uk/aboutus/legaldeposit/websites/ (accessed 23 May 2018).

4. Skinner K, Lippincott S, Speer J and Walters T, Library-as-Publisher: Capacity Building for the Library Publishing Subfield, Journal of Electronic Publishing, 2014, 17(2); DOI: https://doi.org/10.3998/3336451.0017.207 (accessed 23 May 2018).
5. 2001 Budapest Open Access Initiative: http://www.budapestopenaccessinitiative.org/ (accessed 23 May 2018).
6. 2003 Bethesda Statement on Open Access Publishing: http://legacy.earlham.edu/~peters/fos/bethesda.htm (accessed 23 May 2018).
7. Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities https://openaccess.mpg.de/Berlin-Declaration (accessed 23 May 2018).
8. Finch J, Accessibility, sustainability, excellence: how to expand access to research publications, Report of the Working Group on Expanding Access to Published Research Findings, 2012: https://www.acu.ac.uk/research-information-network/finch-report-final (accessed 7 June 2018).
9. Holdren J, Expanding Public Access to the Results of Federally Funded Research, Office of Science and Technology Policy, 2013: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf (accessed 23 May 2018).
10. Wilkinson M D et al., The FAIR Guiding Principles for scientific data management and stewardship, Scientific Data, 2016; DOI: https://doi.org/10.1038/sdata.2016.18 (accessed 23 May 2018).
11. Lavoie B et al., The Evolving Scholarly Record, OCLC Research, 2014: http://www.oclc.org/content/dam/research/publications/library/2014/oclcresearch-evolving-scholarly-record-2014.pdf (accessed 23 May 2018).
12. Lynch C, Stewardship in the 'Age of Algorithms', First Monday, 2017, 22(12); DOI: https://doi.org/10.5210/fm.v22i12.8097 (accessed 23 May 2018).
13. RMap Project: http://rmap-project.info/ (accessed 23 May 2018).
14. Amazon AWS: https://aws.amazon.com/ (accessed 25 April 2018).
15. Microsoft Azure: https://azure.microsoft.com/en-us/?v=18.10 (accessed 23 May 2018).
16. bwFLA – Emulation as a Service, the University of Freiburg: http://eaas.uni-freiburg.de/ (accessed 23 May 2018).

17. Pevner J, 5 March 2018, Yale announces software recovery project, Yale News: https://yaledailynews.com/blog/2018/03/05/yale-announces-software-recovery-project/ (accessed 23 May 2018).

18. Jefferson, 5 March 2018, Andrew W Mellon Foundation Awards Grant to the Internet Archive for Long Tail Journal Preservation, Internet Archive Blogs: https://blog.archive.org/2018/03/05/andrew-w-mellon-foundation-awards-grant-to-the-internet-archive-for-long-tail-journal-preservation/ (accessed 23 May 2018).