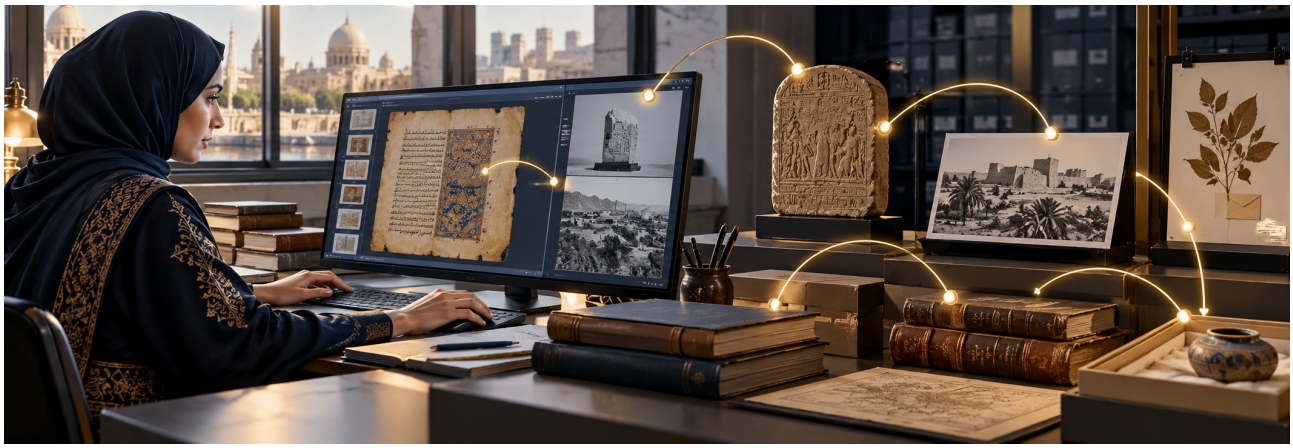
معرّفات (ARK): إدارة أبسط وروابط أدوم للكائنات الرقمية

كايل هوينه، ناتكيران ليدتشوميكانثان، كيرستا شتابلفيلت، عرفان رحمن



الملخص:

يعرض المقال تجربة وحدة البحث الرقمي بجامعة تورونتو سكاربورو في اعتماد معرّفات (ARK) للوصول المستمر إلى الكائنات الرقمية، مع الاستعداد لترحيل منصة (Islandora). ويشرح مقارنة نظم المعرّفات ودوافع الاختيار المرتبطة بالمرونة والكلفة وإدارة البيانات الوصفية. كما يقدم تطبيق (ARKs-Service) لتوليد الأسماء وربطها بالوجهات وتحديثها وحلها وتتبع استخدامها، باستخدام واجهة وجداول بيانات وواجهات برمجية. ويبرز قيمة إنشاء المعرّف مبكرًا وإبقائه مع الكائن خلال الوصف والمعالجة والنشر والترحيل. ويناقش تكامل المعرّفات مع (Islandora) والعناوين المستعارة وبيانات الاستمرارية. ويخلص إلى أن الوصول الدائم يعتمد على الالتزام المؤسسي والصيانة، وأن الأدوات الخفيفة المفتوحة يمكن أن تدعم ذلك دون اعتبار التقنية وحدها ضمانًا للاستمرار. ويؤكد أن اختيار نظام المعرّفات يحتاج تخطيطًا يراعي موارد المؤسسة واحتياجات مجموعاتها.

الكلمات المفتاحية: (ARK)، المعرّفات الدائمة، الكائنات الرقمية، (Islandora)، إدارة البيانات الوصفية

مقدمة:

«الاستمرار مهمة معقدة ومكلفة، ولا تخففها التقنية إلا هنا وهناك» (Kunze, 2003).

مع تزايد الكائنات الرقمية في مجموعات المكتبات، تزداد أهمية ضمان استمرارية الوصول إليها لدى مؤسسات الذاكرة. وتَعِد معظم المستودعات بنوع من الوصول المتواصل. ومن مظاهر الوعد معرّف ثابت، غير دالّ في ذاته، قابل للنقر، أي عنوان (URL) دائم مبني على (ARK) أو (PURL) أو (URN) أو (Handle) أو (DOI) أو (Wikidata QID) أو (ORCID)، وغيرها. وتَعِد هذه المعرّفات باستمرار الوصول وإزالة اللبس في مشهد معلوماتي متزايد التعقيد، وتدعم التزام المكتبيين المهني برعاية مجموعات مجتمعاتهم وبياناتها المميزة.

تتيح وحدة البحث الرقمي (DSU) بجامعة تورونتو سكاربورو كائنات رقمية معقدة متعددة الوسائط عبر تطبيقات ويب عدة، تديرها منصة (Islandora) المفتوحة. وتزايدت طلبات إضافة وظائف الاستشهاد بهذه الكائنات في سياقات أخرى، والاهتمام بنقاط وصول تقرؤها الآلة. وترى الوحدة أن تغذيات (OAI-PMH) وبيانات (IIIF Manifest) تستفيدان من الاعتماد على عنوان واحد معتمد. ومع تطوير نظام (Islandora 2.x) والاستعداد للانتقال، بدا الوقت مناسبًا لإعادة النظر في نظام عناوين بمعرّفات دائمة (PID) . ويعرض المقال منطق اختيار مفتاح المورد الأرشيفي (ARK)، وتطبيق الوحدة لتوليد المعرّفات وربطها وحلها وتتبعها.

اختيار معرّف دائم:

تتوفر أوصاف ممتازة لنظم توليد المعرّفات وصيانتها. يقدم (Leggott et al.) (2016) نظرة إلى أنواعها ووظائفها، و(Barsky) (2020) توصيات كندية، و(Koster) (2020) مقارنة مشابهة، ويحافظ مشروع (Heritage PIDs) على جدول تفصيلي للخصائص (Kotarski, 2021).

تحظى معرّفات (DOI)، المبنية على (Handle)، بمعرفة واسعة في الأوساط الأكاديمية، خصوصًا للمنشورات. وبحسب (DOI.org)، كانت إحدى عشرة جهة تسجيل تقدم خدماتها وقت النشر. ورغم مكانتها معيارًا شائعًا، تجاوزت رسومها ميزانية الوحدة ونطاق مهمتها للمعرّفات الفريدة؛ ويقدم (DataCite) مثالًا على الرسوم. أما (Handle) نفسه فيجذب برسم سنوي رمزي لحفظ البادئات والتسجيل لدى جهة الحل. لكن خصائص (ARK) جعلته في النهاية الأنسب للوحدة.

التعريف بـ(ARK):

يشبه (ARK) نظمًا مثل (DOI) في تمكين إنشاء عنوان دائم وصيانته للوصول الطويل إلى كائن. وظهر في 2001، ويرتبط أكثر بالمجموعات المرقمنة، مع أن مواصفته تدعم كائنات رقمية ومادية ومفهومية عديدة. ويتكون عنوانه من اسم نطاق، ووسم (ARK)، ورقم جهة إسناد الأسماء (NAAN)، الذي يمكن تسجيله مجانًا. وتستطيع المؤسسة تحته تعيين أسماء وأجزاء فرعية وتنويعات. ولا تفرض رسوم على إنشاء المعرّفات.

وتتوفر خدمات مركزية للحل عبر الخدمة المذكورة في المراجع، لكنها ليست شرطًا. ويدعم تحالف (ARK) أكثر من أربعين منظمة، وهو عضو في التحالف الوطني للرعاية الرقمية. وكان سجل (NAAN) المحدث يضم أكثر من تسعمئة مؤسسة عالميًا مخولة بإسناد المعرّفات. ولعل أبرز ما جذب الوحدة خصائص إدارة البيانات الوصفية.

(ARK) وإدارة البيانات الوصفية:

توصي الممارسات الجيدة بإنشاء المعرّف عند ولادة الكائن أو قبلها؛ فنحن نسمي أطفالنا أحيانًا قبل الولادة، ونسمي الأشياء ونشير إليها في مرحلة التصور قبل أن تؤتي ثمارها بوقت طويل.

ويهمنا دعم «الإنشاء المبكر للكائن». فكل كائن تنتجه الوحدة يمر بمراحل تطوير وإغناء، غالبًا بين فرق وذوي مصلحة متعددين، وبسير عمل يعتمد الجداول ويدعمه (ARK) جيدًا. ويمكن توليد المعرّف وإضافته للسجل في أي مرحلة. وفي (Islandora 2.x) يدعم (Islandora Workbench) الإدخال والتحديث مباشرة من الجداول. ثم تتيح عروض (Drupal) وأدوات التسلسل تصدير المعلومات بمرونة، فيسهل تبادل البيانات بين الأدوات واسترجاعها. ويبقى (ARK) مع الكائن عبر الوصف وإعادة التنسيق وضمان الجودة والنشر.

سابقًا، ولّدت الوحدة معرّفات محلية، أو استعانت بمعرّفات نظم مرتبطة، كالأرشيفات. لكنها غير متسقة، ونادرًا ما كان للمعرّف غرض واضح في السجل النهائي. ويدعم (ARK) التعاون في الإنشاء والصيانة لأنه جزء من عنوان يمكن توجيهه وإعادة توجيهه إلى مواقع أخرى. ولتطبيقه مع مراعاة احتياجات مدير البيانات الوصفية، طورت الوحدة (ARKs-Service).

تطبيق (ARKs-Service): التعريف والإعداد:

هو تطبيق (PHP) مستقل بواجهة مبنية على (Noid4PHP) لدانيال بيرثيرو، وهو بدوره تنفيذ (PHP) لمشروع (Noid) الشهير المكتوب بـ(Perl) على يد جون كونز. ويكيف التطبيق عمل أكيو سنسي الذي دمج (MySQL) لتحسين الأداء. ويؤدي تشغيل دليل النشر (playbook) إلى إعداد تطبيق بموقع عام يبسط التوليد والربط والحل، بواجهة وقوالب للمهام الشائعة. ويمكن تغيير الوجهات وتتبع عدد إعادات التوجيه في القاعدة. وتراه الوحدة خطوة نحو حل جاهز لإدارة (ARK)، خصوصًا لفرق التطوير الصغيرة، ويدعم المواصفات القديمة والجديدة.

يوفر التطبيق موقعًا عامًا وواجهة خلفية، يحميها في بيئة الوحدة جدارُ حماية. ويضم العام محللًا يوجه مستخدم عنوان (ARK) إلى الكائن المشار إليه. وبعد التثبيت، تطلب الخطوة الأولى اسم المؤسسة وعنوانها وبيانات مدير النظام.

توليد المعرّفات:

بعد الإعداد، يسجل المستخدم دخوله بالمتصفح وينشئ قاعدة تحتفظ «بمجموعة» من معرّفات (ARK). وبعد تسميتها يختار قالبًا يحدد طريقة التوليد. فلقيم متتابعة من أربعة محارف رقمية ومختلطة مع البادئة الثابتة (sdd)، يختار (sdd.sdede)؛ ولأرقام عشوائية ثلاثية تتوقف بعد القيمة الألف، يختار (.rddd). ويشرح توثيق (NOID) القوالب وخصائصها. وتستخدم الوحدة سلسلة متتابعة واحدة لمعظم مجموعاتها، بالقالب (.zd) الذي يوفر أرقامًا متتابعة بلا حد. والبادئة جزء من المعرّف يجب أن يكون فريدًا عند استضافة قواعد (ARK) متعددة.

بعد إنشاء المجموعة، تتاح وظائف للتوليد والربط والتحديث عبر الجداول. ويدخل المستخدم عدد المعرّفات المطلوب، كمئة مثلًا، فتُنشأ وفق المخطط المختار، ثم يمكن تنزيلها بصيغة (CSV) وفتحها بإكسل أو غيره.

ربط المعرّفات بالعناوين والبيانات:

يمكن بعد التوليد ربط (ARK) بعنوان الكائن في المستودع وأي بيانات وصفية إضافية باستخدام (CSV) آخر. ويوفر التطبيق قالبًا للتنزيل، يمكن إضافة أعمدة لحقول محددة إليه. وحقل (URL) وحده إلزامي تمامًا، مع التوصية بحقول أخرى. ويمكن إلحاق أي عدد من الحقول بالمعرّف. ويمكن أيضًا الربط الفردي من الواجهة.

ويمكن تحديث الربط والبيانات جماعيًا أو فرديًا، وهو مفيد للترحيل. وبعد ربط (URL)، يقود عنوان (ARK) المستخدم مباشرة إلى الكائن. وتتيح إضافة علامة استفهام واحدة أو علامتين إلى آخره عرض لوحة البيانات المرتبطة.

والبيانات المعروضة في المثال لا تكفي لمعيار (ARK)، بل توضح إمكان التطبيق. وسيكون إغناؤها جزءًا من ترحيل الوحدة المقبل، خصوصًا تطوير بيان للاستمرارية، باستخدام مفردات يقترحها (Kunze) (2017)، وتعديل التطبيق لإتاحته عند الطلب.

التقارير:

يمكن في أي وقت تصدير قائمة المعرّفات وبياناتها للمراجعة. ويستطيع المستخدم البحث عن سجل بياناته وتصدير الروابط القائمة. ويسجل النظام عدد مرات استخدام عنوان إعادة التوجيه في عمود (Number of Redirects)، لمعرفة المعرّفات المستخدمة. وتتوفر خصائص الواجهة وتوثيق (REST API) والفيديوهات التفسيرية في ويكي (GitHub) للمشروع، مع الترحيب بالتعديلات وطلبات السحب.

المعرّفات الدائمة في (Islandora):

في (Islandora Legacy)، أي (Islandora 7)، يوفر (Fedora 3.x) معرّف (PID) داخل نطاق أسماء محدد، يظهر جزءًا من عنوان الكائن. وفي نظام الوحدة، حيث تخدم قاعدة (Fedora) نفسها مواقع (Drupal) متعددة، ينتج عن ذلك عناوين مختلفة. ويؤدي العنوانان التاليان مثلًا إلى الكائن نفسه:

(https://gundagunde.digital.utsc.utoronto.ca/islandora/object/gundagunde:17492)

(https://collections.digital.utsc.utoronto.ca/islandora/object/gundagunde:17492)

ويتطلب ذلك صيانة نطاقي (gundagunde) و(collections) الفرعيين، مع إمكان تغير أيهما. وقد يميز (PID)، وهو هنا (gundagunde:17492)، الكائن محليًا في الحالتين، لكن فرادته العالمية غير مضمونة. وواجهت الوحدة أيضًا صعوبات في المعنى الدلالي لأول جزء من المعرّف، إذ يطلب أصحاب المصلحة تغييره. ولذلك كان اعتماد أسماء أقل دلالة ذاتية أولوية مستقبلية.

وفي (Islandora 2.x)، يمكن لـ(ARK) المخزن في بيانات الكائن، وهو كيان (Drupal)، أن يكون أساس اسم مستعار للعنوان يولده (Drupal) آليًا باستخدام (PathAuto). وتحجب أسماء العُقد القائمة على (ARK) رقمَ العقدة الداخلي بمعرّف أكثر دوامًا وتميزًا. ويسمح ربطه بكيان (Drupal) باستخدامه في سياقات أخرى. ويُعرض رابط دائم جديد تحت الكائن، ويُتاح أيضًا في وظائف كتوليد الاستشهادات.

خاتمة:

توجد نظم واستراتيجيات ممتازة كثيرة لتوفير معرّفات وعناوين دائمة للكائنات الرقمية. ووجدت الوحدة أن (ARK) أكثر مرونة وأقل استهلاكًا للموارد للتوليد والربط والحل. ورغم عدم اكتمال الانتقال إلى (Islandora 2.x)، يستخدم بالفعل في سياقات متعددة بالنظام القديم. وتمنح العناوين الدائمة المستخدم قدرًا من الثقة عند الاستشهاد، وتوفر الاستمرارية أثناء الترحيل. ويبني تطبيق الوحدة على برامج مجتمعية مفتوحة ناضجة لإضافة المرونة والواجهة السهلة. ونأمل أن يطور منظومة أدوات خفيفة وميسورة ومرنة لإنشاء المعرّفات وإدارتها.

المراجع:

- Alliance, The ARK. (2021, January 26). ARK implementation best practices. ARK Alliance. Retrieved June 28, 2022, from https://arks.org/about/best-practices/
- Alliance, The ARK. (2022, February 23). Community. ARK Alliance. Retrieved June 28, 2022, from https://arks.org/community/

- Barsky, E. (2020, December 3). Persistent Identifiers in Canada?: White Paper [R]. doi:http://dx.doi.org/10.14288/1.0395014
- Jordan, Mark. Islandora Workbench (2022), GitHub Repository, https://github.com/mjordan/islandora_workbench
- Koster, Lukas (2020) Issue 47 Code4lib journal 20202-02-17 Retrieved from https://journal.code4lib.org/articles/14978
- Kotarski, Rachael, Kirby, Jack, Madden, Frances, Mitchell, Lorna, Padfield, Joseph, Page, Roderic, Palmer, Richard, & Woodburn, Matt. (2021). Developing Identifiers for Heritage Collections (v2.0). Zenodo. https://doi.org/10.5281/zenodo.5205757
- Kunze, J. (2003). Towards electronic persistence using ARK identifiers. Retrieved from https://n2t.net/e/Towards_Electronic_Persistence_Using_ARK_Identifiers.pdf
- Kunze, J., & Rodgers, R. (2008). The ARK Identifier Scheme. UC Office of the President: California Digital Library. Retrieved from https://escholarship.org/uc/item/9p9863nc
- Kunze, J., Calvert, S., DeBarry, J.D., Hanlon, M., Janée, G. and Sweat, S., 2017. Persistence Statements: Describing Digital Stickiness. Data Science Journal, 16, p.39. DOI: http://doi.org/10.5334/dsj-2017-039
- Leggott, Mark, Shearer, Kathleen, Ridsdale, Chantel, Barsky, Eugene, & Baker, David. (2016). Unique Identifiers: Current Landscape and Future Trends. Zenodo. https://doi.org/10.5281/zenodo.557106
- Pathauto. Drupal.org. (2022, April 24). Retrieved June 28, 2022, from https://www.drupal.org/project/pathauto
- Team, D. C. (n.d.). Fee model 2022. DataCite. Retrieved June 28, 2022, from https://datacite.org/feemodel.htm
- N2T resolver: https://n2t.net/