التحقيق والنشر الرقمي: خدمة بحثية جديدة في مكتبة زيورخ

مشروع تجريبي بمكتبة زيورخ المركزية لعرض المجموعات الرقمية بطرائق جديدة

يسكو رايلينغ، شارون روم



الملخص:

يعرض المقال انتقال مكتبة زيورخ المركزية من دعم رقمنة مشروعات الآخرين إلى تنفيذ إصدارين رقميين بوصفهما خدمة بحثية وبنية تحتية للمعرفة. ويتناول مشروعي مراسلات هانس غيورغ وهيرمان نيغلي وإعادة إصدار كتابات جان هيرش، موضحًا مسارات التعرف الآلي على النصوص وتصحيحها وترميزها وربطها بالكشافات. ويناقش مفهوم الإصدار الأرشيفي بوصفه عملًا مستمرًا مفتوحًا للإثراء وإعادة الاستخدام، وأهمية توثيق القرارات التحريرية والتقنية. ويشرح التكامل بين الفهرسة المكتبية وجودة البيانات الوصفية ومتطلبات التحرير الرقمي، مع عرض تحديات الربط بين الفهرس والنصوص الكاملة. كما يصف اختبارات استخدام الذكاء الاصطناعي لإنتاج ترميز (TEI/XML)، ويؤكد تطوير مسارات عمل مستدامة وقابلة للتشغيل البيني تجمع الخبرة المكتبية باحتياجات الباحثين والجمهور العام.

الكلمات المفتاحية: الإصدارات الرقمية، الخدمات البحثية، الإنسانيات الرقمية، البيانات الوصفية، ترميز النصوص

الدخول إلى التحرير الرقمي:

تمتلك مكتبة زيورخ المركزية منذ أكثر من عقدين مركز رقمنة خاصًا وعالي الكفاءة، ترقمن فيه مجموعاتها باستمرار. ومنذ الفترة نفسها تنتج نسخًا رقمية لمشروعات التحقيق والنشر العلمي، وتجمع البيانات الوصفية المصاحبة، وأحيانًا بيانات أخرى خاصة بالمشروع. فقد وصفت ورقمنت خلال ست سنوات أكثر من 23,000 رسالة من يوهان كاسبار لافاتر وإليه، من أجل الإصدار الرقمي لمراسلاته. وأسهمت بكميات أقل في مشروعات أخرى، منها مراسلات كونراد فرديناند ماير وهاينريش بولينغر، ومشروع مراسلات الأطباء في المناطق الناطقة بالألمانية في العصر الحديث المبكر، 1500–1700.

ومنذ مطلع 2024 تنفذ المكتبة بنفسها، وللمرة الأولى، إصدارين رقميين، وتعمل على تطوير وبناء المهارات ومسارات العمل والحلول التقنية في التحرير الرقمي. وبذلك تسلك طريقًا جديدًا وتجرب صيغة جديدة لعرض مجموعاتها للبحث والعلم والجمهور الواسع. وتختبر في الوقت نفسه خدمة بحثية لتوسيع محفظة خدماتها بصورة مبتكرة. ويستهدف المشروعان تطوير مسارات عمل مستدامة وقابلة للتشغيل البيني، تلبي معايير المكتبات ومتطلبات الإنسانيات الرقمية معًا؛ وبناء مسارات ونماذج بيانات خاصة بالمشروع وعامة في آن واحد، يسهل إعادة استخدامها مستقبلًا.

يركز المشروعان على التعرف الآلي على النصوص. ففي مشروع «الإصدار الرقمي لمراسلات هانس غيورغ وهيرمان نيغلي» تُنشر جميع الرسائل التي كتبها ناشر الموسيقى في زيورخ هانس غيورغ نيغلي، 1773–1836، وابنه هيرمان، 1811–1872، أو وُجهت إليهما، مما تحفظه المكتبة. ويبلغ مجموعها نحو 3,800 رسالة كُتبت بين 1790 و1872.

ويولد برنامج (Transkribus) نصًا أوليًا لكل رسالة، ثم يُصحح ويُرمّز وفق معايير (TEI/XML)، ويُربط بكشافات الأعمال والأماكن والأشخاص. وتُنشر الرسائل في النهاية باستخدام (TEI Publisher) على موقع نشر المصادر.

أما المشروع الثاني، «جان هيرش: إعادة إصدار رقمية للكتابات»، فيهدف إلى إتاحة نصوص الفيلسوفة السويسرية المنشورة في مواضع متفرقة. ويُستخدم (Transkribus) هنا أيضًا لإنتاج النص الأولي، ثم تصحيحه وترميزه وفق إرشادات الأرشيف النصي الألماني، وربطه بكشافات الأعمال والأشخاص أو المنظمات.

الإصدارات الأرشيفية عمل مستمر:

تنفذ المكتبة في المشروعين ما يسمى إصدارًا أرشيفيًا، يركز قبل كل شيء على محتوى الوثائق. والغاية تقديم نص قراءة جيد مع كشافات، يلبي اهتمام القارئ العام، ويتيح تحليلات علمية متنوعة، وإعادة الاستخدام والمعالجة اللاحقة في فقه التحقيق المتخصص.

يقدم الإصدار نصًا منضبطًا بترميز أساسي، ولا يسجل جميع الظواهر النصية والمتعلقة بالمحتوى أو يعلق عليها؛ فالترميز الفيلولوجي والمضموني عام عن قصد. ويظل الإصداران منفتحين على أسئلة مختلفة يمكن طرحها عليهما أو تتبعها بواسطتهما. وهما لا يمثلان إصدارين وحيدين مكتملين، كما في كثير من الطبعات المطبوعة، بل عملًا مستمرًا يشكل المرحلة الأولى من المعالجة التحريرية وأساسًا لمشروعات تحقيق مستقبلية.

وهذا يحدد وضعهما الحالي الذي قد يتغير مستقبلًا. وعلى الرغم من أن الإصدار الأرشيفي يتضمن في بنيته طابع الانتقال المستمر الذي يُعد اليوم سمة للأرشيف الرقمي، فإنه يختلف عنه بإضافة وصف أساسي للنص، وتأطير تاريخي للوثائق، وسبل مرنة للوصول إلى المادة المحققة، إلى جانب النص نفسه.

ويتبع الإصدار بناءً مؤلفًا من وحدات، ويظل في جوهره منفتحًا على إعادة الاستخدام والإثراء. وينبع هذا الفهم من طبيعته الرقمية، ويجعل التوثيق بالغ الأهمية؛ إذ يجب تحديث إرشادات الترميز والقرارات التحريرية وجميع المعلومات التقنية وإتاحتها باستمرار، شأن النص المحقق نفسه.

وإعداد الإصدارات الرقمية ليس بحثًا أساسيًا أو خدمة للبحث والعلم فحسب، بل إسهام في بناء البنى التحتية البحثية. وتقدم الإصدارات الأرشيفية لمؤسسات الذاكرة خاصة صيغة رقمية مناسبة لإسهام مهم ومستدام في النقاش العلمي.

من طلب المستفيد إلى خدمة بحثية:

انطلق المشروعان من طلبات مستفيدين. فقد نشأ مشروع نيغلي من بحث مخطط له في معهد علم الموسيقى بجامعة زيورخ. واقترحت جمعية جان هيرش المشروع الآخر، رغبة في إتاحة أعمال الفيلسوفة رقميًا لأهميتها بوصفها صوتًا بارزًا في القرن 20. وتتابع الجمعية تقدم المشروع باهتمام، وتبذل جهودًا ناجحة كبيرة للحصول على تمويل خارجي.

وتجمع إعادة الإصدار الرقمية للمرة الأولى جميع نصوص هيرش المعروفة حاليًا، من أجناس مختلفة: كتب، وفصول في كتب جماعية، ومقالات صحفية ودورية، بل مقابلات ومحاضر جلسات حوار معها. ويُعرض نحو 300 نص في صيغها الأصلية، بالفرنسية أو الألمانية أساسًا، مع الإشارة المنهجية إلى الترجمات اللاحقة إن وجدت.

متطلبات الفهرس المكتبية والتحريرية:

يبدأ مسار العمل في مشروع هيرش من فهرس (Swisscovery). ولأن كثيرًا من عناوينها مسجل أصلًا، تُولّد قائمة العناوين، وهي وثيقة العمل المركزية التي تشمل جميع الأعمال المراد تحقيقها، مباشرة من السجلات الببليوغرافية. لكن كثيرًا من العناوين غير موصوف بصورة كافية، وبياناتها الوصفية ناقصة أو خاطئة، مما يتطلب تنقية السجلات واستكمالها، وأحيانًا إنشائها من جديد.

والهدف ضمان سجل كامل وعالي الجودة لكل عنوان. ويتحقق ذلك عادة بإنشاء تسجيلات تحليلية، مع تنقية سجلات الكتب القائمة عند الحاجة. ولم يكن متوقعًا قبل البدء أن يستغرق فحص البيانات وفهرسة العدد الكبير من العناوين أو إعادة فهرستها كل هذا الوقت.

وتظهر هنا ميزة خاصة للمشروعات التي تنفذ داخل المكتبات: تستطيع المكتبة الاستناد إلى خبرتها لإنتاج بيانات وصفية عالية الجودة في الفهرس، فتنجز العمل المكتبي والتحريري معًا. وتصبح البيانات أساس الخطوات اللاحقة، ويمكن نقلها بسهولة ومباشرة من (Alma) باستخدام رمز مشروع محلي.

يسهل (Swisscovery) التعاون بين المؤسسات في إنشاء السجلات. لكنه يخضع لمتطلبات مكتبية، بما فيها إنشاء مجموعات موضوعية تجمع عناوين محددة. وعند محاولة إنشاء مجموعة لهيرش ظهر قيد: لا يمكن إدراج سوى مقتنيات المنطقة المؤسسية المحلية، التي تحتوي سجلات المكتبة ومقتنياتها، بينما لا يمكن دمج عناوين منطقة الشبكة، المشتركة بين المؤسسات المشاركة. ولأن كثيرًا من عناوين هيرش المراد تحقيقها لا يوجد ضمن مقتنيات المكتبة، فلن يظهر في هذه المجموعة.

وكشف المشروع مسألة أخرى: يمكن ربط واجهات المشروعات بسجلات الفهرس، لكن لا يمكن حتى الآن الإحالة عكسيًا من الفهرس إلى مشروعات التحقيق بسهولة، فضلًا عن الأتمتة. والمطلوب حل تقني يربط السجلات مباشرة بالنصوص الكاملة، ويدمجها أيضًا في مجموعات موضوعية.

ترميز النصوص بالذكاء الاصطناعي:

يعتمد ترميز نصوص هيرش أساسًا على «إرشادات النسخ» التي وضعها الأرشيف النصي الألماني، وهي مواصفة قائمة على معيار (TEI) الدولي للترميز المنظم للنصوص الألمانية التاريخية. وأضيفت إليها إرشادات خاصة بالمشروع.

يُستخدم (Transkribus) مع نموذج (Print M1) للتعرف الآلي وإنتاج النصوص الأولية. ويحقق النموذج نتائج جيدة حتى الآن، بمعدل خطأ في المحارف لا يتجاوز 2.2%. لكن التخطيطات متعددة الأعمدة، الشائعة في الصحف، تمثل تحديًا خاصًا.

توفر وظائف تقسيم الصفحة والوسوم المسبقة في (Transkribus) أساسًا جيدًا لترميز التخطيط وفق (TEI)، مثل وسم الفقرات بـ(<p>) وفواصل الأسطر بـ(<lb>) ووسوم العناوين. لكنها لا تكفي لجميع متطلبات (TEI) المتوافق مع (DTA)، فتستلزم تعديلات وإضافات لاحقة.

ومن هنا نشأت فكرة استخدام أساليب مدعومة بالذكاء الاصطناعي لترميز البنى الشكلية مباشرة بصيغة (DTA) المطلوبة، دون المرور بمخرجات (Transkribus). وبما أن الترميز يتبع قواعد ثابتة، فهو نشاط محكوم بالقواعد لا إبداعي، ويمكن لهذه الطريقة زيادة كفاءة معالجة النصوص الكبيرة.

ويختبر الفريق حاليًا الذكاء الاصطناعي التوليدي للتعرف المنظم على التخطيط والترميز الآلي لعناصر شكلية، مثل الفقرات والكلام المباشر. والهدف توليد وثيقة (XML) كاملة ومستوفية للقواعد بصيغة (TEI) ، باستخدام (ChatGPT) حاليًا. وبعد استقرار الإجراء، تُخطط تطبيقات أخرى، مثل التعرف على الكيانات المسماة وربطها، لتحديد أسماء الأشخاص والهيئات آليًا وربطها بالملف الاستنادي المشترك (GND).

المراجع:

موقع نشر المصادر: https://sources-online.org

Reiling, J. (2020). Bibliotheken ؛ رابط المصدر .Johann Caspar Lavater Online Briefedition, 15 July 2025 :und ihr Beitrag zum Aufbau von Forschungsinfrastrukturen. xviii.ch, 11, 52-62. DOI .10.24894/2673-4419.00005

؛BullingerDigital, 2020–2025 ؛Lukas, W., & Landshuter, S. C. F. Meyer: Digitale Briefedition, 2024 .Stolberg, M. Frühneuzeitliche Ärztebriefe des deutschsprachigen Raums (1500–1700)

يستند المقال إلى محاضرة المؤلفين بالعنوان نفسه في 26 يونيو 2025، في مؤتمر المكتبات التاسع، المنعقد بالتزامن مع (BiblioCon 113) بمدينة بريمن.

Reiling, J., Kreyenbühl, E., & Kohler, D. Digitale Edition der Korrespondenzen von Hans Georg und Hermann Nägeli, 2024. nägeli-briefe.ch.

Sahle, P. (2013). Digitale Editionsformen. Teil 2: Befunde, Theorie und Methodik. Norderstedt, pp. 33–34, 152–154.

.ZBcollections, 2025

.Vorstand der Jeanne Hersch-Gesellschaft. Jeanne Hersch. Die herausragende Genfer Philosophin رابط المصدر

Reiling, J., Kreyenbühl, E., & Rom, S. Jeanne Hersch. Digitale Neuauflage der Schriften, 2025. edition-jeanne-hersch.org.

Bibliotheksinformatik Zentralbibliothek Zürich. ZB Data Map, July 2025. Swisscovery, 2025. رابط المصدر : 2025.

Deutsches Textarchiv. Richtlinien zur Transkription, 2025. رابط المصدر تاريخ الاطلاع لجميع موارد الهوامش الإلكترونية في الأصل: 11 أغسطس 2025.