

Des catalogues de bibliothèques aux projets en humanités numériques : les autorités IdRef font le lien

Les projets de recherche en humanités numériques font un recours croissant à IdRef, qui se déploie ainsi au-delà de son champ documentaire d'origine afin de contribuer à la production de données FAIR.



Les projets de recherche en humanités numériques font un recours croissant à l'utilisation d'IdRef, référentiel d'autorités destiné à décrire des objets documentaires. IdRef se déploie ainsi au-delà de son champ documentaire d'origine et s'incorpore avantageusement aux pratiques de recherche des sciences humaines et sociales, des arts et des lettres, afin de contribuer à la production de données FAIR (*Findable, Accessible, Interoperable, Reusable*).

LES AUTORITÉS DOCUMENTAIRES

L'utilisation d'autorités pour une description qualitative des ressources documentaires des bibliothèques a une longue histoire. La Bibliothèque nationale de France y a eu recours pour mettre en œuvre son Catalogue général. De même, lors de sa création il y a plus de vingt ans, le Sudoc a fait le choix originel d'y recourir.

Les autorités font référence à un nombre restreint d'objets du monde réel dont le rôle est déterminant pour les documents constituant les collections des bibliothèques : Personnes, Collectivités, Familles, Lieux géographiques, Sujets et Œuvres. Plusieurs documents peuvent pointer vers une même autorité : ainsi s'opèrent simultanément économie d'échelle et fédération de contenus.

Pour illustrer ce propos, prenons l'ensemble des romans composant la « Comédie humaine » : ceux-ci sont naturellement liés à la notice de leur unique auteur : « Balzac, Honoré de (1799-1850) ». Entre autre richesse, une notice d'autorité recense les variantes de noms sous une forme normalisée (« H. de Balzac », « Honoré de Balzac »...), ce qui permet la constitution d'index de recherche à la fois riches et précis (ni silencieux, ni bruyants).

Par nature, une autorité a pour vocation l'identification sans ambiguïté de l'objet auquel elle réfère, que ce soit pour distinguer des homonymes ou pour rapprocher des pseudonymes. Au sein d'un référentiel, chaque autorité se distingue donc des autres : « Balzac, Honoré de (1799-1850) » = Lord R'Hoone ≠ Balzac, Henry de (1807-1856).

Rappelons enfin que toutes les données renseignées dans une autorité sont sourcées, établies par la citation de documents et que chaque notice

d'autorité IdRef est dotée d'un identifiant unique et pérenne qui assure la liaison avec les références bibliographiques, en s'appuyant notamment sur le format standard Unimarc.

LA GRANDE FAMILLE DES HUMANITÉS NUMÉRIQUES

Ces caractéristiques rendent évident le fait que les disciplines académiques placées sous la bannière « humanités numériques » sont éligibles à l'utilisation d'autorités. Mieux, elles ont tout intérêt à y recourir. En effet, la documentation, en tant qu'activité spécialisée dans la description et l'accessibilité de ressources et corpus, se conçoit aisément comme une discipline inhérente à ce champ de recherche et de pratiques. Toutes ces disciplines partagent en effet une épistémologie commune fondée sur la citation des sources et le travail sur les objets documentaires, dont la quantité, la variété et la morphologie s'accroissent à l'ère numérique. Par ailleurs, les institutions culturelles telles que les bibliothèques, les musées et les archives élaborent Ides collections qui constituent des sources et données primaires pour les historiens, les chercheurs en littérature et pour toutes les disciplines des sciences humaines..

Même si la reconnaissance de cette situation n'est pas encore universellement partagée, elle est pourtant porteuse d'importants bénéfices collectifs.

LA TRAJECTOIRE D'IDREF

Après que le Sudoc a fait le choix de cataloguer sous contrôle d'autorités, l'Abes a été chargée en 2005 de développer Calames, application spécifique au signalement des Manuscrits et Archives des bibliothèques de l'ESR, qui se devait d'intégrer cette bonne pratique. Pour ce faire, les autorités du Sudoc, jusqu'alors confinées en coulisses, se sont vues d'un seul coup exposées sur le « World Wide Web », donnant ainsi naissance à IdRef, application de réutilisation et d'édition mutualisée de ces autorités.

Ouvertes à Calames, les portes le sont depuis à quiconque souhaite réutiliser et, mieux encore, contribuer à l'enrichissement des référentiels d'autorité dont IdRef est le vecteur. C'est ce qu'a fait

Persée dès 2014¹, date à laquelle la plateforme de ressources numérisées en SHS a migré sa base d'Auteurs et intégré dans son *workflow* le liage et la création d'autorités IdRef.

Au-delà du Sudoc, au-delà de l'Abes, un seuil est franchi et se profile alors la trajectoire idéale d'une utilisation d'IdRef, depuis les grands catalogues de bibliothèques en passant par les archives institutionnelles et les bibliothèques numériques jusqu'aux projets de recherche en humanités numériques.

LES SOLLICITATIONS DES HUMANITÉS NUMÉRIQUES

En la matière, l'Abes commence à disposer d'un portefeuille de sollicitations assez garni : que ce soit spontanément ou par l'intermédiaire d'appels à projet (au premier rang desquels CollEx-Persée), leur nombre ne faiblit pas. Illustrations avec trois projets qui montrent une utilisation et une prise en main simples et efficaces des autorités IdRef à des fins de constitution d'index et d'identification fiable.

- Le projet SIPROJURIS², Système d'information des professeurs de droit (1804-1950) a « idrefisé » son corpus de professeurs.
- Le projet PRELIB³, projet de recherche en littérature de langue bretonne, porté par le Centre de recherche bretonne et celtique (CRBC) et soutenu par la Maison des sciences de l'homme de Bretagne, a « idrefisé » l'ensemble des personnes, comme cela se voit dans l'interface d'IdRef :

Recherche sur Miossec, Yves idref 027031357

MIOSSEC, YVES (1907-2001)

CRBC Bibliographie dans PRELIB (Projet de recherche en littérature de langue bretonne)

7 références PRELIB : Tout voir dans PRELIB

(affichage limité aux 30 premières références)

- PRELIB 1916 Drest ar mor bras, Miossec Yves, Brud Nevez - Emgleo Breiz 1981
- PRELIB 1915 Eur veaj e Stadoù Unanet an Amerik, Miossec Yves, 1996
- PRELIB 1917 Lohr, kasoni ha poan, Miossec Yves, Brud Nevez - Emgleo Breiz 1983
- PRELIB 1918 Mond da Israel, Miossec Yves, Brud Nevez - Emgleo Breiz 1991
- PRELIB 1914 O haloupad war heñchou Breiz-Izel, Miossec Yves, 1987
- PRELIB 2353 Tammoù ennoennoù euz eun amzer griz, Miossec Yves, Brud Nevez - Emgleo Breiz 1994
- PRELIB 2254 Tammoù ennoennoù euz eun amzer griz, Miossec Yves, Emgleo Breiz - Fondation culturelle bretonne 2013

- Le projet Marc Michel Rey⁴, projet de publication de la correspondance de ce libraire du XVIII^e siècle porté par l'IHRIM (Institut d'histoire des représentations et des idées dans les modernités), a « idrefisé » les expéditeurs et destinataires, donnant lieu à la génération d'environ 70 notices à partir des informations recueillies par les chercheurs, ce qui a permis de construire un index dédié au projet, comme par exemple :

INDEX DES PERSONNES

Accarias de Serionne, Jacques  073754684

Journaliste né à Die, rédacteur du Journal de commerce et d'agriculture (1759-1762).

L'intérêt croissant du recours au référentiel IdRef par les chercheurs et ingénieurs impliqués dans ce type de projet s'explique aisément par la nature même des données d'autorité. En effet, à l'aide des technologies déployées par IdRef en vue de leur « FAIRisation », chaque notice est dotée d'un identifiant unique et pérenne, pivot indispensable à l'interopérabilité entre différents systèmes d'information ; indexée, elle est aisément trouvable, requêtable et réutilisable, notamment du fait de sa mise à disposition sous licence ouverte dans les formats propres au web sémantique (RDF). De plus, les données peuvent être enrichies et sérialisées selon les besoins spécifiques de chaque projet.

UN ACCOMPAGNEMENT RENFORCÉ

Cependant, données et outillage ne seraient pas si efficaces sans l'accompagnement que les chercheurs peuvent trouver auprès de l'Abes ou, localement, auprès des bibliothécaires et professionnels de l'information ayant développé une expertise de la structuration et de l'interopérabilité des données.

En éclaircur, l'Abes explore cette voie depuis 2019 en relation étroite avec des archéologues de la Maison de l'Orient et de la Méditerranée, pourvoyeurs de données hétérogènes, et des historiens, experts en ontologie et sémantisation des savoirs du Laboratoire de recherche historique Rhône Alpes dans le cadre du projet ANR HisArc-RDF⁵ (pour HISToire et ARChéologie en triplets RDF) qui vise la création d'une chaîne de traitement menant à la production de données de recherche de qualité FAIR. Est récolté ainsi le fruit de l'expérience et de l'investissement dans les technologies du Web sémantique, qui s'incarne dans `data.idref`⁶, et pour ce projet même, dans un point d'accès *sparql* expérimental alliant accompagnement et exposition en RDF des jeux de données homogénéisées⁷, où se déploient dialogue, échange et acculturation réciproques et interdisciplinaires.

Continuité épistémologique, mutualisation des outils, des services et des savoir-faire, interdisciplinarité de l'effort heuristique : le chemin est encore long mais l'horizon est là qui nous guide pour que les communautés formant la grande famille des humanités numériques concourent à l'élaboration de leur grand corpus unifié de connaissances liées et ouvertes.

FRANÇOIS MISTRAL

Responsable IdRef-Autorités
mistralf@abes.fr

[1] <http://info.persee.fr/lalignement-des-autorites-persee-au-referentiel-idref>

[2] <http://siprojuris.symogih.org>

[3] <https://mshb.huma-num.fr/prelib>

[4] <http://rey.huma-num.fr/presentation>

[5] <https://anr.fr/fr/lanr/engagements/la-science-ouverte/les-projets-laureats-de-lappel-flash-science-ouverte/projet-hisarc-rdf>

[6] <https://data.idref.fr>

[7] <https://dataforhumanities.abes.fr/sparql>