

La curation, un enjeu pour la gestion des données numériques

La gestion des jeux de données numériques est devenue un enjeu incontournable de la science ouverte. La curation de ces données intègre des tâches de sélection, de vérification, de normalisation, de structuration ou encore d'enrichissement, indispensables pour publier des données réexploitables. Dans ce domaine, les bibliothécaires ont pour atout leur expérience collective et partagée de tout ou partie de ces tâches. Avec l'ouverture des catalogues et la diffusion de leurs métadonnées, ils sont déjà devenus des gestionnaires de données potentiellement utilisables par la communauté des chercheurs.

Parmi ces gisements de données, les autorités – que la bibliothéconomie moderne préfère appeler « entités » – sont les divas. Gardons à l'esprit qu'elles sont de plus en plus nombreuses et chouchoutées : Ces données *de référence* constituent en effet des pivots facilitant l'interopérabilité des systèmes. C'est bien l'usage souhaité qui détermine les traitements à effectuer. Il peut s'agir de produire des alignements vers ces données pour augmenter la visibilité d'un *corpus*, ou bien de les enrichir pour mieux les partager, ou encore tout bonnement de les créer pour que tout système puisse s'y raccrocher. Dans cette optique, les catalogueurs changent de pratique, en intégrant à leurs tâches quotidiennes des « opérations qualité » ciblées. Pour les accompagner, l'Abes investit dans des outils et services adaptés à ces nouvelles réalités.

PAPRIKA, UNE INTERFACE POUR VOIR ET AGIR AUTREMENT

Ouverte en février 2019, l'application Paprika (paprika.idref.fr) propose une autre manière de visualiser et de produire des données dans le Sudoc et dans IdRef. À partir d'un nom et d'un prénom, l'utilisateur voit sur un même écran l'ensemble des autorités et des points d'accès correspondants, et accède à des vues construites jusqu'alors inaccessibles : une personne dans le contexte d'un document particulier, un auteur à travers la somme de tous les documents auxquels il a contribué. *In fine*, Paprika permet de créer et de corriger des liens et de créer des notices d'autorité sans passer par l'interface de catalogage WinIBW. Sa simplicité et son caractère ludique et visuel en font un outil pédagogique exploité d'ailleurs par certains coordinateurs Sudoc pour « dédramatiser » le catalogage et impulser une prise de conscience commune de l'importance des liens aux autorités.

QUALINKA, UN PROGRAMME D'INTELLIGENCE ARTIFICIELLE AU SERVICE DU SIGNALLEMENT

Qualinka est un service d'intelligence artificielle capable de produire et d'évaluer des liens entre des points d'accès et des autorités personne. Disponible à partir de l'application Paprika, il offre un diagnostic automatisé de la qualité des liens. Sans se substituer au catalogueur qui reste seul habilité à modifier les données, Qualinka a pour vocation de lui faciliter la tâche

en focalisant son attention sur les cas problématiques. En dehors de Paprika, Qualinka est utilisé en mode automatique pour effectuer des alignements en masse.

ALGOCLC2 : LA FRBRISATION DU SUDOC, VERS LA TRANSITION BIBLIOGRAPHIQUE

L'expérimentation « Sudoc-FRBR » permet la préfiguration d'un nouveau type d'entité du modèle LRM : l'œuvre. L'algorithme de calcul baptisé « algoclc2 » produit précisément des pré-notices d'œuvres issues du regroupement de notices bibliographiques. À l'heure actuelle, les catalogueurs n'ont pas la possibilité de créer ou modifier ces pré-notices. En revanche, ils peuvent influencer sur la qualité des regroupements, en corrigeant des erreurs de catalogage (fautes de frappe, liens erronés aux personnes ou aux titres uniformes) ou bien en enrichissant les notices trop pauvres (titre incomplet, absence de zones de liens).

Ce tour d'horizon rapide montre une partie des moyens mis en œuvre par l'Abes pour soutenir les professionnels du réseau Sudoc en matière de curation des données d'autorité. Après des premiers pas prometteurs, reste à encourager chacun (professionnel, équipe, établissement) à se saisir de ces dispositifs.

ALINE LE PROVOST

Service Autorités et Référentiels, Abes
le-provost@abes.fr

