Le projet collaboratif Mistara réactualise les problématiques d'alignement et de qualité des données en écriture non latine tout en les confrontant aux nouveaux outils de l'Abes et de la BnF.

Mistara : au croisement des enjeux de la transition bibliographique et de la recherche sur l'onomastique - ou étude des noms - arabo-musulmane

RETOUR D'EXPÉRIENCE

La mistara, tablette de réglure utilisée pour tracer les lignes d'écritures dans les manuscrits du monde islamique, est la métaphore parfaite pour un projet centré autour de l'alignement des noms de personnes en écriture arabe. Le projet Mistara[1], financé par le GIS CollEx-Persée de 2018 à 2021, a été porté par la Bulac (Bibliothèque universitaire des langues et civilisations) en partenariat avec le service commun de la documentation d'Aix-Marseille Université, la section arabe de l'Institut de recherche sur l'histoire des textes (CNRS), l'Institut dominicain d'études orientales du Caire, la bibliothèque de l'Institut du monde arabe et la Maison méditerranéenne des sciences de l'Homme, avec le soutien des agences bibliographiques nationales. La vertu première de ce type de financement a été d'encourager une démarche collaborative. Elle a permis de réactualiser et de traiter des problématiques déjà connues d'alignement et de qualité des données en écriture non latine, tout en les confrontant aux nouveaux outils élaborés par l'Abes et la BnF dans le cadre de la transition bibliographique. Le projet visait à utiliser des référentiels tiers issus de la recherche : la base AlKindi de l'Institut dominicain d'études orientales [Ideo] et l'*Onomasticon Arabicum* de l'Institut de recherche et d'histoire des textes [IRHT] pour enrichir IdRef.

Le projet croise les enjeux de la transition bibliographique et de la recherche sur l'onomastique arabo-musulmane, en cherchant d'une part à évaluer les adaptations nécessaires à la modélisation et à la manipulation des formats d'entités personne, à la gestion d'entités multilingues et multiécritures et aux spécificités de la décomposition du nom dans le monde arabo-musulman, et d'autre part à améliorer la gestion des référentiels produits par la recherche sur les questions onomastiques du monde arabo-musulman pour assurer leur interopérabilité et permettre l'enrichissement des outils catalographiques.

UN TRAVAIL COOPÉRATIF ENTRE PARTENAIRES ISSUS DE DIFFÉRENTS HORIZONS

Les objectifs du projet ont été atteints en bonne partie. Les données des différents partenaires ont permis d'enrichir les données du Sudoc et de tester leur traitement avec différents outils d'alignement ou de correction des liens entre notices : Bibliostratus, Openrefine, Qualinka/Paprika. L'hétérogénéité des données de départ a tout d'abord nécessité un traitement humain avant de pouvoir utiliser les outils existants. Un rapport détaillé sur le bilan d'usage de ces outils pour des données en écriture arabe et des données décrivant des entités physiques a été élaboré avec l'Abes. Les consignes de traitement des entités en caractères arabes ont pu être précisées, les bases d'exemples enrichies pour préciser et homogénéiser la prise en compte des pratiques d'écriture et la richesse onomastique du monde islamique : ces informations seront mises à disposition du réseau de l'Abes sous la forme d'un guide complémentaire pour le catalogage des autorités arabo-musulmanes.

La dimension coopérative du travail autour de l'alignement des référentiels s'est avérée particulièrement riche. Une dynamique nouvelle a pu se créer entre des partenaires dans et hors du champ de l'enseignement supérieur, et des opportunités intéressantes ont surgi. Une collaboration est entamée entre l'Ideo et l'Abes qui devrait permettre aux établissements qui utilisent l'environnement *ILS-DIAMOND* [nativement LRM] de produire leurs données dans IdRef et, à terme, d'enrichir rétrospectivement le référentiel avec les notices produites jusqu'à ce jour. Ce sera un enrichissement essentiel du point de vue de l'onomastique arabe qui profitera également aux études de la patristique latine et orientale couvertes par les établissements travaillant avec *ILS-DIAMOND.* Au-delà du projet Mistara, la collaboration entre les partenaires du projet et l'Abes se poursuit donc !

Cécile Gobbo
Responsable du Pôle Flux et données de la Bulac
cecile.gobbo@bulac.fr

[1] L'expérience liée au projet et ses apports ont fait l'objet d'un carnet de recherche sur https://mistara.hypotheses.org/193

© Maxime Ruscio/BULAC

Détail d'une mistara ottomane, tablette de réglure (collections de la Bulac).