

Lillanalytics : un service documentaire en appui au pilotage de la recherche

Lauréate en 2017 d'une initiative d'excellence I-Site, l'Université de Lille oriente sa politique de recherche autour de 4 hubs thématiques, dont l'un « Cultures, sociétés et pratiques en mutation » dédié aux Sciences humaines et sociales (SHS).

La stratégie de recherche de l'Université s'appuie depuis 10 ans sur une culture forte de la bibliométrie fondée sur la plateforme Lillometrics, portée en coopération avec le CHU de Lille. Au sein de l'Université, le portage de Lillometrics est initialement assuré par la Direction de la valorisation de la recherche, avec l'appui du Service Commun de Documentation (SCD). Le contexte interne est celui d'une forte coopération entre les deux équipes et d'une identification des responsabilités respectives sur les dossiers communs.

Sur ce terreau favorable émerge en 2020 le besoin de caractériser les domaines d'expertise et les collaborations variées de l'Université, avec un accent particulier sur les SHS, peu couvertes par la bibliométrie. Le projet vise à concevoir une base de connaissances requêtable, en mettant à profit des sources de données diversifiées et en mobilisant des technologies de fouille de texte. La Direction de la valorisation demande au SCD d'assurer le portage de ce projet et de réaliser une étude de faisabilité, avec l'appui d'un financement I-Site.

Le SCD engage d'abord de premières expérimentations, avec la conception d'une *proof of concept* (POC) sur son corpus de thèses de doctorat. Ce POC vise à extraire des remerciements des thèses toutes les entités nommées relatives à une collaboration avec des partenaires. Cette première expérience formatrice confirme les difficultés posées par l'usage du *text mining* : la diversité des questions que l'on souhaite poser et la qualité relative des sources de données rendent difficile la création d'une base de connaissances fondée uniquement sur de la fouille de texte.

Le projet Lillanalytics connaît alors une seconde phase. Nous décidons de nous appuyer plus concrètement sur les besoins

et usages de la gouvernance de l'Université, en développant sans attendre un service de fourniture d'analyses. La démarche consiste à exploiter les métadonnées d'articles, ouvrages et chapitres d'ouvrage, des projets de recherche, les pages professionnelles des chercheurs, les thèses de doctorat en préparation ou soutenues ainsi que le texte des rapports d'auto-évaluation HCERES. Les analyses produites permettent d'affiner les besoins de la gouvernance et de rendre le projet immédiatement visible et utile. En 2024, Lillanalytics soutient par exemple la préparation de DemoCIS, la réponse de l'Université de Lille à l'appel à manifestation d'intérêt SHS.

En partant des commandes de la gouvernance et des directions, il devient plus facile de définir des priorités pour la complétude et la qualité des données. Un travail de recensement est mené en parallèle pour décrire les sources, leurs caractéristiques et leur couverture. Il permet notamment de repérer les responsables des sources, d'identifier des besoins d'amélioration des sources exploitées et d'engager des collaborations entre services autour de l'enjeu de la mise en qualité.

Les premières réussites du projet conduisent à son renouvellement, puis à sa pérennisation sous la forme d'une mission en 2025. La partie universitaire de Lillometrics et Lillanalytics sont réunis au sein d'une nouvelle cellule d'appui au pilotage de la recherche, positionnée au SCD. Cette approche permet de mutualiser les compétences documentaires, bibliométriques et d'analyse de données pour un meilleur service de la gouvernance, des Directions centrales et des unités de recherche. Cette pérennisation marque la reconnaissance par tous les acteurs de l'importance d'une mission articulant fourniture d'analyses de

la production scientifique et fiabilisation des sources de données.

À partir de cette démarche peu technologique, plusieurs perspectives s'ouvrent désormais pour prolonger le travail engagé :

- Explorer l'intégration des technologies d'intelligence artificielle dans nos dispositifs
- Déployer une stratégie locale autour des identifiants, en cohérence avec la stratégie nationale française
- Mutualiser les efforts pour ajouter les briques technologiques manquantes, tout en valorisant l'existant local
- Constituer en interne un collectif d'acteurs dédié à la fiabilisation des données relatives à la recherche.

Ces enjeux passent par la construction d'une culture professionnelle commune et partagée entre les services acteurs de la gouvernance des données sur la recherche.

ELSA DREVON

chargée de mission Lillanalytics,
Université de Lille
elsa.drevon@univ-lille.fr

MADELEINE GÉROUDET

responsable du Département Services à la
recherche et aux chercheurs, Université de Lille
marie-madeleine.geroudet@univ-lille.fr

- Consulter la présentation :
<https://doi.org/10.5281/zenodo.15704251>
- Consulter la vidéo de l'intervention :
<https://vimeo.com/1088421549>



Credit photo Adobe Stock - Tetiana, généré à l'aide de l'IA.