

# Recommandation de citation : étendre les références aux textes plein

## Contexte

La recommandation de citation est une tâche qui consiste à proposer des citations pour un texte donné. En raison de l'explosion du volume des travaux scientifiques publiés ces dernières années d'une part, et de la nécessité de citer les publications les plus appropriées lors de la rédaction de textes scientifiques d'autre part, la recommandation de citation est devenue un sujet de recherche important. Malgré les efforts entrepris par la communauté scientifique, un verrou au développement de systèmes de recommandation de citation reste l'absence de collection de test standardisée pour en évaluer leur efficacité. L'objectif de ce TER est de participer à la création d'une telle ressource en mettant au point des outils pour la récupération automatisée d'articles scientifiques en accès libre.

## Objectifs

- Étudier la collection de test pour la recommandation de citation actuellement en cours de création (e.g. format des données, sources utilisées, méthodologie de construction)
- Étendre la collection avec les textes plein (pdfs) des articles en accès libre. Il faudra pour cela implémenter un scraper.
- Évaluer le niveau d'efficacité des méthodes existantes sur la collection de test.

**Nombre d'étudiants** : 1/2 étudiants

**Encadrants** : Florian Boudin, Béatrice Daille

**Lieu** : LS2N