

Contribution à l'exploration d'un corpus de documents textuels pour l'analyse du discours public

Contexte

Le projet DISTIC vise à analyser le discours et la prise de décision publique, en particulier dans le contexte de crises épidémiologiques ou environnementales (pollution au mercure en Guyane, accidents industriels...). Il rassemble des chercheurs en sciences humaines et sociales, en sciences de l'information et en informatique. Le projet dispose de corpus documentaires hétérogènes (mail, rapports, articles, supports de communication) dont l'analyse doit permettre de faire émerger les logiques de prises de décision publique ou les engagements citoyens.

Travail du stage

Dans le cadre de ce projet, les missions qui seront confiées à l'étudiant·e sont les suivantes :

- Contribution à la définition d'une méthodologie pour l'analyse de ces données en collaboration avec les autres membres du projet
- Implémentation de la nomenclature et des outils retenus
- Indexation des documents et de leurs métadonnées
- Extraction d'informations (OCR, entités nommées)
- Visualisation (création de tableaux de bord et analyse statistique)

Profil du stagiaire

- Étudiant·e de M1 informatique
- Expérience en analyse de données (statistique, texte)
- Connaissance du langage Python
- Capacité à travailler en équipe

Lieu du stage et durée du stage

- Laboratoire L3i, La Rochelle Université
- 10 à 12 semaines, gratification légale

Contact

Pour toutes questions ou candidatures, merci de contacter :

- Cyrille Suire : cyrille.suire@univ-lr.fr
- Guillaume Bernard : guillaume.bernard@univ-lr.fr

Merci d'inclure dans vos candidatures, CV, relevés de notes de L3 et 1^{er} semestre de M1, et résumé court des projets informatiques que vous souhaitez valoriser.