

# Sujet de PSTL :

# Scikit-network : une bibliothèque de manipulation de grands graphes en python

**Encadrants :** Maximilien Danisch

## Objectifs

Les moteurs de recherche, la conception de médicament et la gestion du trafic ne sont que quelques exemples d'applications qui reposent sur l'analyse de graphe. Les graphes concernés sont de plus en plus massifs, atteignant parfois plusieurs trillions de liens.

L'objectif du stage est de participer au développement de scikit-network, une bibliothèque émergente de manipulation de grands graphes en python. En particulier, de réaliser des fonctions pour calculer une approximation de pagerank efficacement (la méthode push).

## Tâches à réaliser

- Réaliser des implémentations efficaces de fonctions en python pour la bibliothèque scikit-network.
- Rédiger une documentation et faire intégrer les fonctions à la bibliothèque.
- Comprendre et restituer l'analyse théorique en temps et en mémoire des algorithmes implémentés.
- Réaliser des expériences en comparant les différentes méthodes.

## Prérequis

- Bonne connaissance de python et de son optimisation (e.g. cython).
- Connaissance basique du langage C.
- Un intérêt pour l'algorithmique des grands graphes.

## Références

- <https://scikit-network.readthedocs.io/en/latest/>
- <https://sites.google.com/view/limass/software>s
- <https://papers-gamma.link/paper/174> (méthode push détaillée page 6)