

# Université Nice Sophia Antipolis

## M2 & SI 5 option AL

Le barème est donné à titre indicatif, mais il peut être modifié.
---

Février 2020 - Durée 2H00

Dans vos réponses, soyez synthétique, précis, appuyez vos propos sur vos lectures, votre expérience, ...

**Question :** La société Poly4All développe une nouvelle application. Elle a besoin d'y intégrer des outils logiciels qui facilitent le traitement de textes en utilisant de la linguistique et de l'apprentissage automatique. Elle a trouvé le projet NLP-Square en open-source qui répond à ces exigences fonctionnelles. Ce projet a déjà 7 ans, est disponible sur github, sous gestion de version, avec un suivi par ticket élaboré. Le projet utilise la technologie docker, est en partie écrit en java, C et python. La société vous a mandaté pour conduire une étude sur ce projet avant de l'intégrer ou non dans sa nouvelle application.

Elle veut des informations sur

- 1- la qualité du projet NLP-Square, en particulièrement relativement aux risques d'erreurs en fonction des différents modules/outils proposés;
- 2- les dépendances du projet relativement aux logiciels externes utilisés;
- 3- la dynamique d'évolution du projet relativement à la communauté, la correction des erreurs, l'intégration de nouveaux algorithmes de traitements des textes;

Nous attendons de vous d'expliciter comment vous allez procéder et SURTOUT pas d'anticiper sur les résultats. Vous prendrez le soin, quand cela est adéquate, de préciser votre positionnement relativement aux travaux présentés dans l'article (*par exemple*, réutilisation, perspectives, extensions).

Article support : Nagappan, Nachiappan, et al. "Mining Metrics to Predict Component Failures." *Proceedings - International Conference on Software Engineering*, 2006, doi:10.1145/1134285.1134349.