

## PROPOSITION DE STAGE - API TELEGRAM

### Contexte

Initialement lancée en tant que messagerie sécurisée, l'application Telegram est progressivement devenue la principale source d'information dans des nombreux pays.

Toutefois, de nos jours, certains médias diffusent parfois des informations erronées ou des messages de propagande faisant qu'il s'avère compliqué de distinguer l'information authentique des fake news.

En outre, des organismes générateurs d'informations semblant indépendants les uns des autres, peuvent promouvoir les mêmes renseignements et créant ainsi l'illusion d'un grand nombre de sources.

Partant de là, grâce au traitement de texte en langage naturel et à l'API de Telegram – laquelle permet de suivre ces chaînes efficacement – une opportunité importante se présente : celle de créer des outils pour s'informer correctement et vérifier les informations avant de les transmettre aux autres.

Il y existe une plateforme similaire, [www.buster.ai](http://www.buster.ai), dont un point discutable est qu'elle accepte quelques sources en tant qu'a priori fiable. La plateforme [propaganda.qcri.org](http://propaganda.qcri.org) permet de détecter la présence des procédés standards de la propagande, elle semble avoir été abandonné depuis 2020. Enfin, s'il existe d'autres projets semblables comme [hoaxbuster.com](http://hoaxbuster.com), [factcheck.org](http://factcheck.org), [snopes.com/](http://snopes.com/), [observers.france24.com](http://observers.france24.com), [afp.com](http://afp.com) et [signal-arnaques.com](http://signal-arnaques.com), mais ces derniers ne sont automatisés.

### Profil et compétences bienvenues

- étudiant en Bac+5 en informatique ;
- expérience ou intérêt pour : Python, Go, Node.js, l'intégration à une API existante, le déploiement des applications multi-langages, l'implémentation des solutions répondant aux problèmes réels de la vie courante, et ce projet en général ;
- rigueur, créativité, ouverture d'esprit.

## **Mission et quelques détails techniques**

Le travail consistera en plusieurs tâches parmi les suivantes :

- analyse des messages d'un grand nombre de sources d'information ;
- extraction, avec OpenAI, de l'idée générale d'un texte d'actualité ;
- détection des similitudes entre les médias non liés officiellement ;
- construction d'un graphe des médias dépendants ;
- création d'une API publique analysant une information transmise ;
- analyse des messages en temps réel, dont les limites sont imposées par des tiers (OpenAI) – la scalabilité étant une partie très importante du projet (implémenté en Python) ;
- analyse sur demande d'un message gérant en simultanéité un certain nombre d'utilisateurs (implémenté en Go ou Node.js) ;
- utilisation du fil d'actualité des médias pour l'analyse en temps réel ;
- tests des différentes approches et de leurs effets sur la performance et le coût de la maintenance.

La réponse de l'API contiendra

- l'idée principale d'une information ;
- un résumé de l'idéologie promue par la chaîne ;
- une liste des médias similaires ;
- une liste de médias sur la même thématique, mais diffusant des idées différentes.

Les nouveaux messages des chaînes seront analysés et stockés. Il faudra expérimenter afin de trouver des schémas efficaces pour le stockage et l'analyse. L'API publique permettra de consulter ces données stockées (ou, éventuellement, utilisera un modèle efficace en temps réel).

Les bases de données utilisées seront : MongoDB et Elasticsearch.

Les processus en arrière plan seront déclenchés par la file des tâches et les flux d'événements.

## **Détails pratiques**

La structure d'accueil : l'association Ouverture du cœur (2017) concentrée sur des projets Humanitaires et Solidaires.

Localisation Locaux associatifs : AERI, 57, rue Étienne Marcel 93100 Montreuil.

Durée : 5 mois, 140 heures par mois, avec la date de début le 3 avril 2023.

La gratification : 4,05 € / heure

Président de l'Association Ouverture du Cœur : Mr Kopera Tony : 0753338738,  
Mail : [ch4nch4n@live.fr](mailto:ch4nch4n@live.fr).

Responsable Technique Mr Seren Guillaume : [guillaumeseren@gmail.com](mailto:guillaumeseren@gmail.com)

Une équipe de 8 informaticiens : Mr Pierre de Lacroix, Mr Sébastien Kurt, etc.