

# Deep learning based semantic search applied to recruitment platforms

<b>Contacts</b>	Aziz AMAL, ALTEN R&D, <a href="mailto:aziz.amal@alten.com">aziz.amal@alten.com</a> Zoltan MIKLOS, IRISA, <a href="mailto:zoltan.miklos@univ-rennes1.fr">zoltan.miklos@univ-rennes1.fr</a>
<b>Localisation</b>	ALTEN Innovation Center, Rennes
<b>Durée</b>	6 mois
<b>Mots clés</b>	Natural Language Processing, Deep Learning, Reinforcement learning

## Contexte

Le Responsable RH a pour mission de sélectionner les candidats pertinents pour les postes à pourvoir, proposer des parcours attractifs aux salariés et favoriser un environnement de travail épanouissant pour les collaborateurs. Cependant, certaines tâches chronophages, telles que la présélection (premier tri) de curriculum vitae, le détournent de son cœur de métier. A titre d'exemple, Google recevrait chaque semaine 75 000 CV.

Cette problématique de présélection de candidats et d'évaluation de la pertinence de candidature par rapport à un besoin peut être assimilée à la conception d'un moteur de recherche ayant comme requêtes des offres de travaux et proposant les profils professionnels associés.

Outre la mise en correspondance des offres de travaux et des profils professionnels, fonctionnalité principale du dispositif, il peut être utilisé aussi, comme un **outil d'aide à la validation des offres de travaux**. Ceci lui permet d'assister le Responsable RH à parfaire son offre en améliorant son efficacité par la restitution, pour validation, des concepts (et leurs importances ou forces) contenus dans l'offre. La performance de l'offre pourra aussi se mesurer sur un jeu de test choisi.

Le projet vise à répondre aux besoins internes de l'entreprise tels que la mobilité et la gestion de carrière et externes tel que le recrutement de candidat.

## Problématique

L'enjeu principal est d'évaluer l'utilisation des techniques de « deep learning » dans la caractérisation et la mise en correspondance des profils professionnels et des offres de poste.

Ce projet est une démarche concurrente à plusieurs projets ALTEN avec la même finalité mais basés sur des méthodes classiques. L'objectif sera aussi d'avoir une approche « comparée ».

## Questions scientifiques et techniques

L'objectif des études à mener est de répondre aux incertitudes suivantes :

1. Concevoir un/des modèles permettant l'extraction de connaissances à partir des profils professionnels et des offres.
2. Définir les critères de choix du corpus de profils destiné à l'apprentissage automatique.
3. Concevoir et mettre en œuvre une méthode d'évaluation de la pertinence du/des modèles
4. Concevoir un/des modèles d'apprentissage interactif en utilisant les techniques du deep learning et du reinforcement learning (intégration de l'évaluation humaine suite à une recherche)

## Cadre du stage

Le stage est localisé à Rennes, et est rémunéré et pourrait se poursuivre, en fonction des résultats obtenus, par une thèse CIFRE.

## Bibliographie

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, Jeffrey Dean, 2013, Distributed Representations of Words and Phrases and their Compositionality, CoRR

Quoc V. Le, Tomas Mikolov, 2014, Distributed Representations of Sentences and Documents, CoRR

Jeffrey Pennington, Richard Socher, and Christopher D. Manning. 2014. GloVe: Global Vectors for Word Representation

Alexis Conneau, Holger Schwenk, Loïc Barrault, Yann Lecun, 2017, Very Deep Convolutional Networks for Text Classification, EACL 2017

TensorFlow, <https://www.tensorflow.org/>

Keras, <https://keras.io/>