

Stage COEXEL

Développement d'un module de recommandations appliqué à une solution logicielle de veille (Big Data)

➤ *Données générales*

Démarrage : Courant 2020

Durée : Stage (4-6 mois)

Lieu de travail : Toulon (+ déplacements/séjours à Paris)

Rémunération : Base légale + TR

Divers : Prolongation possible du stage en CDD si satisfaction mutuelle

➤ *Description de l'entreprise*

La société COEXEL est une jeune entreprise spécialisée en veille stratégique et intelligence économique. Elle appuie son développement sur la fourniture de prestations de services à forte valeur ajoutée dans les domaines de la veille et de l'intelligence économique.

Parallèlement, la société a développé un outil web de gestion électronique de l'information, appelé *mytwip*, pour répondre aux attentes de ses clients en matière de veille stratégique. La vision à long terme pour COEXEL quant au développement de cette solution est de devenir leader sur le marché du web sémantique et des outils d'intelligence économique.

➤ *Problématique et contexte*

Les outils de veille technologique et stratégique permettent de délivrer des services de recherches d'information et de notifications de données ciblées, que ce soit en direct ou en temps réel. Ces données ciblées correspondent à des évolutions technologiques visibles sur le Web pour lequel un expert du domaine souhaite rester au courant de la concurrence ou des usages dans son périmètre.

La difficulté pour ces outils de veille est de devoir traiter d'une part les données avec à la fois la multitude de domaines d'expertise pour répondre à la demande des experts, acquérir et gérer un grand volume de données à récupérer sur le Web, analyser le contenu des informations pour en ressortir de la pertinence. Et d'autre part, gérer le profil des experts sur leurs usages de recherche, d'interactions avec la plateforme de veille, mais également les connaissances de l'expert sur son environnement, comme sa propre base de connaissances ou un réseau d'experts.

La société Coexel se positionne dans ce domaine de la veille technologique & stratégique en proposant la plateforme *MyTwip* dédiée à ne nombreux domaines d'expertise, avec un moteur de recherche dédié, intégrant des analyses sémantiques basées sur une ontologie pour classifier automatiquement les informations par domaines d'expertise, du traitement de textes pour identifier des signaux faibles pour détecter ces évolutions technologiques pertinentes, ou de l'extraction de connaissances pour relier les concepts liés à une information.

➤ **Sujet de stage**

Afin de mieux intégrer l'expert dans le processus de veille, nous envisageons d'intégrer le profil utilisateur, l'expert, au sein même de l'environnement de recherche à différents niveaux. Concrètement, le projet consiste en la mise en place d'un système de recommandations se basant sur les précédentes interactions de l'utilisateur. Celles-ci, de nature hétérogène (recherches, lecture détaillée, documents, notations, suppressions, commentaires, validations manuelles) permettent de caractériser le besoin de l'expert et, par conséquent, de lui suggérer de nouveaux documents ou nouvelles sources susceptibles de l'intéresser.

Il s'agit donc de trouver une pondération pertinente de ces différentes interactions afin de définir précisément le profil de l'expert et, une fois celui-ci spécifié, d'identifier sources et documents les plus proches de lui. Il y a donc une notion de distance à affiner entre les différentes entités mises en jeu (documents, sources, profils d'experts) et entre plusieurs instances d'une même entité (on peut ainsi envisager d'étudier la proximité entre deux profils d'experts P1 et P2 afin de recommander à P1 les documents validés par P2 dans le cas où P1 et P2 seraient suffisamment similaires)

Ainsi, la croisée de ces différents critères produit un système complexe dont le mélange particulier a pour but de produire des informations pertinentes aussi bien par recherche à la volée qu'en temps-réel. Cette combinaison subtile n'est pas traitée dans la littérature, en effet nous comptons pouvoir produire des recommandations pertinentes de manière efficace en temps réel avec des profils multidimensionnels en se basant à la fois sur des historiques d'événements et sur des graphes de connaissances ou un réseau social.

Ce défi à relever repose sur certains verrous que nous devons soulever :

- Définir une mesure de pertinence de recommandations reposant sur un profil utilisateur riche, reposant sur son historique étendu (données explicites & implicites), son réseau social et son graphe de connaissances ;
- Définir un système optimisé pour la recherche d'information et la recommandation de veille technologique, combinant temps-réel et traitements lourds pour des milliers d'expert.

➤ **Profil**

Formation :

De formation informatique (Bac+5 minimum), vous avez une expérience significative dans le développement et avez une expérience dans le déploiement et la gestion de projets ou souhaitez évoluer vers ce type de poste.

Connaissances requises :

- Maîtrise générale des Bases de Données relationnelles.
- Maîtrise générale du langage Python et du format JSON
- Connaissance des concepts de recherche d'information, métadonnées et web-sémantique.
- Capacité à gérer efficacement un workflow conséquent

Connaissances appréciables :

- Expérience de l'outil de gestion de projet Redmine
- Expérience de la BDD No-SQL Elasticsearch.

Qualités nécessaires :

- Motivation
- Rigueur
- Autonomie

➤ **Candidature**

Envoyer par email et en pdf les documents suivants :

- CV
- Lettre de motivation ciblée sur le sujet
- Lettres de recommandation
- Relevés de note + Enseignements suivis en M1 et M2

➤ **Contact**

Vincent BOISARD

PDG

Tel : 06 22 05 40 89

Email : vincent.boisard@coexel.com