

Automatic ontology learning and semantic search in HR domain

Contacts	Aziz AMAL, ALLEN R&D, aziz.amal@alten.com Zoltan MIKLOS, IRISA, zoltan.miklos@univ-rennes1.fr
Localisation	ALLEN Innovation Center, Rennes
Durée	6 mois
Mots clés	Natural Language Processing, Machine Learning, Ontology Learning, Semantic Web

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

Elaboration d'un modèle d'extraction et de représentation des connaissances des profils professionnels et des offres permettant d'effectuer des recherches sémantiques sur ces derniers. Ce projet s'inscrit dans la continuité des travaux de 3 projets réalisés par ALLEN (3 * 6 mois) et qui ont eu pour objectif d'exploiter le TALN (NLP), le web sémantique et l'apprentissage automatique pour extraire et hiérarchiser des concepts permettant la représentation des profils professionnels et des offres en vue de les mettre en correspondance.

Ce projet est une démarche concurrente à plusieurs projets ALLEN 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. Comment identifier automatiquement les concepts présents dans un profil ou une offre ? Et comment extraire les relations inter-concepts ?
2. Comment évaluer la performance du dispositif par rapport à ses enjeux ?

3. Concevoir et mettre en œuvre une méthode d'évaluation des distances entre concepts ? et concevoir un modèle de représentation tenant compte des relations inter-concepts et des distances entre ces derniers ?
4. Concevoir et mettre en œuvre une méthode de recherche (mise en correspondance entre les offres de poste et les profils professionnels)

Cadre du stage

Le stage est localisé à Rennes, et est rémunéré et pourrait se poursuivre, selon l'avancement, par une thèse CIFRE.

Bibliographie

Lin He, Han-Qing Hou, 2008, Research on Semi-Automatic Construction of Domain Ontology based on Machine Learning and Clustering Technique, Intelligent Information Technology Application Workshops 2008

Qiuxia song, jin liu, xiaofeng wang, jin wang, 2014, A novel automatic ontology construction method based on web, Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP) 2014

Ryan Ribeiro de Azevedo, Fred Freitas, Rodrigo G. C. Rocha, José Antônio Alves de Menezes, Cleyton Mário de Oliveira Rodrigues, Gabriel de F. P. e Silva, 2014, An approach for learning and construction of expressive ontology from text in natural language, Web Intelligence (WI) and Intelligent Agent Technologies (IAT) 2014

Kessler, R., G. Lapalme, & É. Tondo, 2016, Génération d'une ontologie dans le domaine des ressources humaines, CORIA 2016

Kiruparan Balachandran, Surangika Ranathunga, 2016, Domain-Specific Term Extraction for Concept Identification in Ontology Construction, Web Intelligence 2016

Rosario Girardi, 2016, An Analysis of Main Solutions for the Automatic Construction of Ontologies from Text, Web Intelligence 2016

Hannah Bast, Björn Buchhold, Elmar Haussmann, 2016, Semantic Search on Text and Knowledge Bases

Apache Stanbol - Welcome to Apache Stanbol (<https://stanbol.apache.org/>)

Haro Rafa and Morales Antonio Pérez, 2016, Structuring Medical Records with Apache Stanbol