




mnemotix
Donnons du sens à vos données

RECONNAISSANCE AUTOMATIQUE DE COMPÉTENCES DANS LES TEXTES DESCRIPTIFS DES FORMATIONS

Rome-OpenEmploi

Référence	Open Emploi_Reconnaissance automatique des compétences_2.pdf
Statut du document	 Cette œuvre est mise à disposition selon les termes de la Licence Creative Commons Attribution - Partage dans les mêmes conditions 4.0 International
Date d'émission	16-11-2020
Émetteur	Pierre-René Lhérisson (Mnemosyne)
Destinataires	Olivier Coreau
Version	1.0



Introduction

Pour illustrer l'automatisation de la reconnaissance des compétences dans les offres de formation, nous avons appliqué nos méthodes de reconnaissance de concepts sur des textes qui décrivent ces formations.

1. Données CARIF

Nous sommes partis de la description de la formation suivante :

BTS techniques physiques pour l'industrie et le laboratoire

<https://www.espace-competences.org/Formation/Action/1447950>

2. Données Éduscol

Nous sommes partis de la description de la formation :

Bac Pro Aéronautique option avionique

<https://eduscol.education.fr/sti/formations/bac-pro/bac-pro-aeronautique#referentiel>

3. ROME

Nous utilisons les compétences qui sont dans le ROME comme base de données

4. Travail réalisé

Nous avons cherché à reconnaître de manière automatique les compétences qui seront acquises après avoir suivi des formations professionnelles. Nous avons abordé ce sujet, comme un problème de reconnaissance de concepts dans du texte.

En effet, les compétences sont des **SKOS concept** organisés sous forme de thésaurus et elles sont répertoriées dans un serveur de recherche (Elastic Search). Nous avons utilisé la percolation, qui contrairement au comportement habituel d'un serveur de recherche qui est d'indexer des documents, sauvegarde des requêtes. Le service de percolation permet de passer en paramètre un document qui sera ensuite comparé aux requêtes sauvegardées au préalable.



Nous avons donc dans un premier temps sauvegardé les labels qui définissent les compétences du ROME sous forme de requêtes. Nous avons utilisé la syntaxe de requête de Lucene. Dans un second temps nous avons lancé le service de percolation en pensant au moteur de recherche le texte descriptif de la formation. Le service nous répond en nous envoyant les compétences qu'il retrouve dans le texte.

5. Résultat

Nous présentons dans cette partie les résultats obtenus pour chaque texte descriptif de formation :

BTS techniques physiques pour l'industrie et le laboratoire
Transmettre les techniques du métier ou mener des actions de sensibilisation au métier
Veiller au respect de la loi Informatique et Libertés et du RGPD dans l'entreprise, gérer la liste des traitements de données à caractère personnel, faire l'interface avec la Commission Nationale de l'Informatique et des Libertés - CNIL
Qualité
Hisser, fixer et régler un câble et ses accessoires en fonction de la température, de la distance entre les poteaux et des données de plan de piquetage
Concevoir le schéma directeur à partir des orientations fixées par la direction et des besoins des services utilisateurs et superviser les modalités de mise en oeuvre
Planifier les interventions des équipes de piste et la mise à disposition du matériel et des équipements pour la touchée (temps d'immobilisation au sol)

Bac Pro Aéronautique option avionique
Transmettre les techniques du métier ou mener des actions de sensibilisation au métier
Hisser, fixer et régler un câble et ses accessoires en fonction de la température, de la distance entre les poteaux et des données de plan de piquetage
Contrôler le dossier de vol, le plan de vol, les données techniques de l'appareil et vérifier le niveau de carburant et les conditions météorologiques
Veiller à la sécurité, à la salubrité et à la tranquillité publique et mettre en oeuvre des opérations de secours en cas d'incident ou de catastrophe naturelle



Vérifier le fonctionnement des appareils de projection, la qualité du son et de l'image et effectuer les réglages des projecteurs et des chaînes sonores

Planifier les interventions des équipes de piste et la mise à disposition du matériel et des équipements pour la touchée (temps d'immobilisation au sol)

6. Conclusion

Ce document fait état d'un travail réalisé dans le cadre du projet Open Emploi. Le but a été de présenter une méthode de reconnaissance des compétences dans des textes descriptifs de formation.

Nous présentons les résultats obtenus par le système. Tous les concepts présentés ont un score de correspondance inférieur à 2% (le meilleur score est 100%). Ceci s'explique par la difficulté que le système a à trouver des intitulés de compétence dans les textes.

Les résultats obtenus font état des améliorations à apporter au système de reconnaissance de concepts. Nous savons empiriquement que le système fonctionne mieux quand le texte de description est succinct. En recherche d'information dans l'état de l'art, ils utilisent les résumés (*abstracts*) des textes, pour notre cas d'usage nous pouvons faire un résumé automatique du texte avant de faire la reconnaissance de concepts. Nous pouvons aussi attendre des organismes de formations des fichiers structurés avec une liste des compétences définies. Nous pouvons restreindre les compétences aux nœuds de compétences dans un index spécifique. Les nœuds de compétences comptent moins de tokens. En effet les compétences "longues" ont tendance à créer des requêtes longues qui vont correspondre à tout texte.