

Stage en Traitement Automatique des Langues H/F :
SUJET : DETECTION AUTOMATIQUE MULTILINGUE DE SOUS-PARTIES DE MOTS (MORPHEMES)

Avec plus de 90% de son CA à l'international, et plus de 100 millions d'utilisateurs dans le monde, Vision Objects est un éditeur de logiciels leader mondial sur le marché des interfaces homme-machine basées sur la reconnaissance d'écriture manuscrite.

Disponible dans plus de 85 langues, ses produits concernent les marchés de la mobilité (saisie de texte, prise de notes, ...), de l'éducation (apprentissage de l'écriture, des mathématiques, de la géométrie, ...) de l'entreprise (prise de notes et traitement de formulaires), et de l'automobile (saisie de texte à partir d'une surface tactile, interaction avec GPS).

Vision Objects est une entreprise d'innovation et de hautes technologies. Le cœur de sa technologie MyScript est diffusé sous forme de kit de développement logiciel, de « Cloud service », de composants à intégrer ou sous forme d'applications prêtes à l'emploi.

Le moteur de reconnaissance de Vision Objects se classe régulièrement aux premières places des compétitions scientifiques internationales (cf., par exemple, ICDAR). Dans le cadre de sa forte croissance, Vision Objects (Nantes, France) est à la recherche d'un:

Stagiaire Ingénieur Informaticien en Traitement Automatique des Langues (TAL)

Dans l'équipe *Ressources Linguistiques*, vous serez amené à travailler sur la mission suivante :

SUJET : DETECTION AUTOMATIQUE MULTILINGUE DE SOUS-PARTIES DE MOTS (MORPHEMES)

Les langues fortement agglutinantes posent des problèmes spécifiques en modélisation statistique des langues, notamment le très grand nombre d'unités lexicales possibles. Une approche est de découper ce qui est couramment appelé mot en unités plus petites.

Le stage consiste à étudier et implémenter des algorithmes non supervisés (sans exemples de découpage dans la langue cible) de découpage de mots en morphèmes. Il s'agit ensuite d'appliquer ces algorithmes dans un contexte fortement multilingue car notre solution est disponible en 64 langues.

Une suite possible de ce travail sera l'étude de techniques également non supervisées et multilingues d'analyse grammaticale (PoS tagging).

Stage basé à Nantes, démarrage dès que possible en 2013 ou 2014.

Rigoureux, dynamique et d'un relationnel facile, vous saurez rapidement vous intégrer au sein des équipes.

Au sein de VisionObjects, vous travaillerez sur des technologies à la pointe de la recherche et pourrez identifier les applications directes et concrètes de votre travail.

Un ancien stagiaire, promotion ECN 2010-2013, a ainsi contribué à l'élaboration de l'application Android MyScript Calculator: une calculatrice manuscrite utilisant la reconnaissance d'équations de VisionObjects. Les applications réalisées et publiées par VisionObjects à partir du travail de ce stage ont depuis totalisé plus de 10 millions de téléchargements.

Vous pouvez trouver gratuitement MyScript Calculator sur les stores Android et iOS.

Contact : job@visionobjects.com

VisionObjects

3 rue de la rainière - 44339 Nantes cedex 3 – France

Tel: +33 (0)2 28 01 49 50 - Fax: +33 (0)2 40 25 89 20

www.visionobjects.com

SA au Capital de 1 380 560 Euros – 419 963 806 RCS Nantes

TVA intracommunautaire : FR 43 419 963 806