

Offre d'emploi : EXTRACTION D'INDICES SONORES DANS LES TEXTES ET LES IMAGES, IDENTIFICATION DE CONCEPTS ET RECHERCHE D'INFORMATION MULTIMODALE

Type de contrat : CDD	Type de poste : Post-doctorant	Formation requise : doctorat
Durée du contrat : 18 mois	Rémunération : Catégorie A en fonction de l'expérience, sur la base de la grille de l'enseignement supérieur et de la recherche, ingénieur d'études (~ 3600 € Brut/mois)	Prise de fonction : 01/02/2020

Le LIRIS recherche un post-doc (H / F)

DESCRIPTION DU POSTE

Le (la) candidat.e exercera ses missions au sein du laboratoire LIRIS (équipe Imagine) sous la responsabilité hiérarchique de Véronique Eglin du LIRIS (veronique.eglin@insa-lyon.fr) et de Mylène Pardoën (mylene.pardoen@msh-lse.fr) de la MSH-LSE.

Organisation du travail : 37h

Temps complet

Lieu de travail : Laboratoire LIRIS, INSA Lyon, Bâtiment Joseph Fourier, 69621 Villeurbanne Cedex

Les activités scientifiques du LIRIS sont structurées en 6 pôles de compétences, dont le pôle vision intelligente et reconnaissance visuelle dont l'équipe Imagine est membre (21 permanents, 30 doctorants). Les thématiques principales sont les suivantes: Construction d'algorithmes visant la compréhension de données multimédia (images, vidéo, documents numérique, scènes 3D) et se déclinant en termes d'acquisition/reconstruction, d'indexation, de modélisation, de classification ou de reconnaissance automatique du contenu (objets, actions, concepts).

Missions principales : L'objet principal de la mission vise la mise en place d'un système de recherche multimodale, dynamique et interactif qui intègre différentes sources de données (données sources issues du patrimoine numérisé des bibliothèques et archives de la Bibliothèque Nationale de France) et autorise l'association de **requêtes à la fois textuelles et de caractéristiques visuelles** pour permettre une recherche d'images par le texte, ou par le contenu visuel. Le système sera conçu selon une conception reposant sur un bouclage de pertinence avec l'utilisateur, permettant le raffinement de la recherche et de l'exploration des sources.

PROFIL RECHERCHÉ

Savoirs : indexation et classification d'images, reconnaissance de formes, techniques d'apprentissage, recherche d'information par le contenu, word-spotting, wrd-embedding (BERT ou ELMO)

Savoir-faire : maîtrise des environnements de développement, en particulier de la programmation C++ et usage de la librairie OpenCV, utilisation de bibliothèques de deep-learning (Pytorch ou Tensorflow), réalisation de prototype fonctionnel

Savoir être : aisance dans le cadre d'un travail de groupe pluridisciplinaire, ouverture et communication avec des spécialistes des sciences humaines et sociales, autonomie et capacité à prendre des initiatives

CANDIDATURES

Renseignements sur le poste : veronique.eglin@insa-lyon.fr

Envoi des candidatures : lettre de motivation + recommandation + CV exclusivement par e-mail avant le 25/01/2020 minuit à veronique.eglin@insa-lyon.fr

DÉTAIL DU PROFIL

Le laboratoire LIRIS recrute un **postdoc de 18 mois** sur la **recherche d'information multimodale**. Ce poste est destiné à rejoindre un programme de prématuration du CNRS SYMTESENS (SYstème MulTimodal d'Extraction de données SENSorielles hétérogènes) qui couvre des thématiques autour de l'Archéologie du paysage sensoriel, les données hétérogènes, l'analyse sensorielle multimodale, le wordspotting, les approches sémantiques de la fouille de texte, l'apprentissage automatique. Le projet vise plus largement à industrialiser la production de paysages sensoriels telle qu'elle a été réalisée jusqu'ici dans le cadre du projet Bretez (reconstruction sonore du quartier du Châtelet dans le Paris du 19ème).

L'objectif est de concevoir et mettre au point une **méthodologie complète de recherche d'information** (de la méthode au moteur de recherche) dans les documents numérisés des archives (textes imprimés OCRisés ou non, textes manuscrits numérisés de faible qualité) des périodes 18ème-19ème siècle relatives à la ville de Paris. Cette méthodologie sera initialement fondée sur la **construction d'un dictionnaire terminologique conçu selon des techniques de word-embedding** basées sur une modélisation du langage et des modèles attentionnels incorporant le contexte (de type BERT ou ELMO).

Habituellement découplées, les recherches d'information en mode «plein texte» et «image» permettent d'extraire des sources documentaires des fragments informationnels pertinents. Chacun des domaines de recherche (texte et image) dispose de son lot de techniques qui ont fait leur preuve dans des domaines où l'accès au contenu ne pouvait se faire que par une seule modalité d'interrogation. La recherche dans les images de textes manuscrits par exemple est habituellement résolue par les techniques alternatives à l'OCR (techniques dites de word-spotting dont l'équipe Imagine du LIRIS est spécialiste) lorsque la transcription est trop complexe à obtenir. On procède ainsi pour un corpus d'images à des approches de ce type basées sur des requêtes soit images, soit textuelles soit encore symboliques qui aboutissent à la détection de fragments d'images similaires à la requête. L'idée ici est de pousser plus loin la recherche par le contenu en créant une véritable **interaction texte/image** (dans un espace de représentation joint liant les mots dans leur représentation textuelle et image) et qui permettra une interrogation par requêtes multimodales. L'espace de représentation joint sera bâti à partir de séquences visuelles (images de mots, séquences de descripteurs) liées à un **lexique de termes sonores** initialement conçu pour le projet à partir de la modélisation de concepts sonores à différents niveaux : indices directs (ex : « le chien aboie ») à indirects (ex : « le forgeron travaillait dans la rue») et les ambiances sonores associées (ex : « la rue était bruyante » ou encore « on entendait souffler le vent »). Ce **dictionnaire de termes** sera rendu possible grâce à une sélection d'un **sous-corpus transcrit et annoté** issu des données réelles de terrain. Il servira ensuite à l'ensemble des apprentissages nécessaires à la construction d'un espace de représentation joint liant les mots dans leur représentation textuelle et image. Un second corpus d'images de textes également annoté sera ensuite exploité à des fins de tests, selon une **interrogation par requêtes textuelles** pour la recherche d'information.

Ce travail donnera lieu à la mise au point d'un outil interactif de **prévisualisation des données** permettant aux utilisateurs de naviguer et de consulter des pages en profondeur et d'aide au raffinement de la requête. Les compétences techniques sur la question du **bouclage de pertinence** et de hiérarchisation (ranking) des réponses du système d'interrogation expertes seront abordées pour assurer une interaction entre l'utilisateur et le système. Un prototype d'interrogation constituera le dernier livrable du postdoc et sera soumis à l'évaluation par les praticiens sur les lieux où sont conservées les sources documentaires.