

CONTRAT POST-DOCTORAL / ASSISTANT DE PROJET SUR LA THEMATIQUE : « INTELLIGENCE ARTIFICIELLE POUR LES LANGUES PEU DOTÉES »

Durée : 12 mois

Type de contrat : Contrat Post-doctoral (Décret n° 2021-1450 du 4 novembre 2021 relatif au contrat post doctoral de droit public)

1 poste à pourvoir

L'Inalco est un établissement public unique au cœur du Nouveau quartier latin. Créé en 1795, il est le seul établissement public d'enseignement supérieur et de recherche au monde à proposer une offre de formation en langues et sciences humaines et sociales aussi riche et reconnue en France comme à l'international avec plus de 100 langues et civilisations enseignées.

L'Inalco enseigne auprès de 9 000 étudiants en formation initiale. L'Institut compte plus de 500 personnels.

Le recrutement s'effectue dans le cadre de la Chaire de Professeur Junior "Intelligence artificielle pour les langues rares ou peu dotées" dont l'Inalco est lauréat.

Le projet vise à **relever les défis posés par les langues disposant de ressources numériques limitées**. De nombreuses langues à travers le monde ne disposent pas des données linguistiques et des outils informatiques nécessaires pour bénéficier des avancées récentes en traitement automatique des langues (TAL) et en intelligence artificielle (IA). Ce projet **explore des méthodes pour développer des technologies basées sur l'IA capables de traiter, de faire de la recherche et d'enseigner ces langues, en mettant l'accent sur la rareté des données, la diversité linguistique et l'interopérabilité multilingue**. L'objectif est de concevoir des modèles robustes pouvant soutenir diverses applications linguistiques, allant de l'analyse textuelle à la traduction automatique, tout en préservant la diversité et l'accessibilité des langues.

CONTEXTE ET OBJECTIFS DU POSTE

Une partie de ce projet est consacrée au **traitement de la parole pour les langues peu dotées**, avec un accent particulier sur la **reconnaissance automatique de la parole (ASR)** et les **systèmes de synthèse vocale (TTS)**. Les technologies de la parole nécessitent de grands ensembles de données annotées, souvent indisponibles pour ces langues, en particulier dans des contextes dialectaux variés. La recherche porte sur les stratégies de collecte de données, les techniques d'augmentation et les modèles d'IA capables de fonctionner efficacement dans des environnements multilingues et multidialectaux. En développant des méthodologies pour l'entraînement de modèles de parole avec des ressources limitées, ce projet contribue à l'avancement des applications de l'IA pour les langues orales, améliorant ainsi leur utilisation dans les domaines éducatifs, culturels et technologiques.

Nous recherchons un chercheur postdoctoral spécialisé en traitement automatique de la parole, avec des méthodes issues du traitement automatique des langues (TAL) et de l'intelligence artificielle (IA), appliquées aux langues peu dotées.

Le candidat retenu travaillera dans le cadre de la Chaire Professeur Junior

"Intelligence artificielle pour les langues peu dotées", afin de faire progresser la recherche sur le traitement des données orales, avec un accent particulier sur les défis multidialectaux et les scénarios de code-switching.

Le chercheur recruté travaillera en étroite collaboration avec le titulaire de la Chaire "Intelligence artificielle pour les langues rares ou peu dotées" ainsi qu'avec l'équipe ERTIM (Équipe de Recherche Textes, Informatique, Multilinguisme) de l'INALCO. Site web de l'ERTIM : <https://www.inalco.fr/ertim>.

Le chercheur postdoctoral se concentrera sur :

- Le traitement de la parole pour les langues rares ou peu dotées, en relevant les défis liés aux systèmes de reconnaissance automatique de la parole (ASR) et de synthèse vocale (TTS).
- Le développement de modèles robustes capables de gérer la variation dialectale et les continuums linguistiques.
- L'exploration de méthodes de collecte, d'annotation et d'augmentation des données pour améliorer la performance des modèles dans des contextes à faibles ressources.
- L'expérimentation de techniques d'IA pour optimiser la technologie vocale dans des environnements présentant une grande variabilité linguistique et peu de ressources disponibles.
- La possibilité de développer des outils pédagogiques pour les langues sous-représentées, basés sur des technologies d'IA.

MISSIONS ET ACTIVITÉS

Les activités scientifiques du ou de la post-doctorant/e concerneront :

- Participation, soutien et collaboration aux activités de la Chaire "Intelligence artificielle pour les langues peu dotées".
- Organisation d'un événement scientifique (journée d'étude en mai 2026, ateliers) en lien avec le projet de recherche et les activités de la Chaire.
- Participation régulière aux activités scientifiques de l'ERTIM.
- Organisation d'un terrain de recherche pour la collecte de données orales dans une langue peu dotée (optionnel).
- Rédaction et publication d'un ou deux articles dans des revues scientifiques à comité de lecture.

COMPÉTENCES ATTENDUES

- Doctorat en linguistique computationnelle, TAL, apprentissage automatique ou domaine connexe, obtenu après 2022.
- Expérience approfondie en reconnaissance automatique de la parole (ASR), synthèse vocale (TTS) et traitement de la parole en général.
- Familiarité avec la modélisation des langues peu dotées et les défis liés à la diversité linguistique.
- Expérience dans le travail avec une langue peu dotée.
- Compétences en programmation en Python et maîtrise des frameworks de deep learning (ex. PyTorch, modèles de langage LLMs).
- Excellentes compétences en rédaction scientifique.

INFORMATION S RH

- **Type de contrat** : Post-doctoral d'une durée de 12 mois
- **Rémunération brute** : 2 500€ mensuel
- **Temps plein** : 38h45 hebdomadaire

-
- **54 jours de congés payés annuels** dont 2 périodes de fermeture obligatoire (3 semaines en été et 1 semaine à Noël) ;
 - Tous les postes de l'Inalco sont **ouverts aux personnes en situation de handicap** ;
 - **Télétravail** : Des jours de télétravail réguliers ou ponctuels, selon les nécessités de service, sans condition d'ancienneté ;
 - **Localisation du poste** : ERTIM, INALCO, 2 rue de Lille, 75007, Paris ;
 - **Date de prise de poste souhaitée : à compter du 16 juin 2025.**

Le **dossier de candidature** composé de :

- Une copie du diplôme de doctorat ou une attestation de réussite.
- Une lettre de motivation expliquant leur intérêt et leur adéquation avec le poste.
- Un CV détaillé accompagné de la liste des publications.
- Un résumé du projet de recherche (deux pages maximum).

Le dossier de candidature est à retourner sous forme électronique au plus tard le **15 mai 2025** aux adresses suivantes :

Mme Valentina Fedchenko, Chaire de Professeur junior :
valentina.fedchenko@inalco.fr et copie à drh-recrutement@inalco.fr
