

N° Galaxie : 4248

N° de l'emploi : 774

Nature de l'emploi : **CPJ**

Section CNU : 27, 07

APPEL À CANDIDATURES

Chaire de professeur junior :

Intelligence Artificielle pour les langues rares ou peu dotées

PROFIL DE POSTE

L'Institut National des langues et civilisations orientales lance un appel à candidatures pour un poste de Chaire de Professeur Junior sur le thème : « Intelligence Artificielle pour les langues rares ou peu dotées ».

Il s'agit d'un recrutement sur un contrat à durée déterminée (5 ans) avec une charge d'enseignement réduite, permettant à terme, et après évaluation, d'être titularisé sur un poste de professeur des universités fonctionnaire.

Prise de fonction au plus tard **le 31 décembre 2024**.

Le dispositif se déroule en deux temps :

1. Un appel à candidatures sur un projet de recherche et d'enseignement qui permettra de conclure un contrat à durée déterminé en tant qu'agent contractuel de la fonction publique, pour une durée de 5 ans ;
2. A l'issue de la période de contrat, une commission de titularisation se réunira pour évaluer la valeur scientifique et l'aptitude professionnelle permettant une titularisation dans le corps des professeurs des universités. Au moment de la titularisation, sera exigée l'obtention d'une HDR. La publication d'au moins 6 articles dans des revues à comité de lecture est attendue aux termes des 5 années de pré-titularisation. Il est aussi attendu, pendant la période de pré-titularisation, que la personne titulaire de la Chaire participe aux projets collectifs initiés au sein de l'établissement et ait une activité importante de valorisation du programme (conférences, séminaires, activités de diffusion de la recherche...).

Contrat de chaire de professeur junior : contrat à durée déterminée de droit public d'une durée de 5 ans

Section CNU : 27, 07

Mots-clés : Intelligence Artificielle ; Traitement Automatique des Langues ; Langues peu dotées ; Numérisation ; Reconnaissance de la parole

Rémunération : Indice 735 de la fonction publique française (environ 3680€ bruts par mois)

Montant du financement associé : La chaire de professeur junior bénéficiera d'un forfait pour frais de recherche de 200 000 € alloué par l'Agence nationale de la recherche (ANR) pour la période de 5 ans. Le programme de la Chaire pourra par ailleurs bénéficier d'un complément de financement provenant d'autres partenaires nationaux et internationaux. Celui-ci pourra permettre de financer des postes de chargés de projets, doctorants contractuels, post-doctorants, la participation à des conférences, l'accès à des bases de données, des équipements, le financement de souscription à des services (algorithmes ou données), etc.

Conditions à remplir par les candidats : Les candidats doivent être titulaires d'un doctorat, et avoir un dossier de recherche bien établi montrant une activité de recherche dans des domaines pertinents pour le profil du poste (voir ci-dessous), comme des publications dans des revues internationales de premier ordre, des implications scientifiques dans des projets de recherche, etc.

Le candidat recruté devra démontrer sa maîtrise de l'une des langues de son aire de spécialisation ainsi que sa capacité à assurer des enseignements en anglais.

Projet de recherche

L'Institut National des Langues et Civilisations Orientales (Inalco) est un grand établissement public d'enseignement supérieur et de recherche reconnu pour ses travaux portant sur les langues, les cultures et les sociétés du monde au travers de disciplines de Sciences Humaines et Sociales telles que la linguistique, la sociologie, l'économie, la géographie, l'histoire, les relations internationales, la science politique, les études littéraires et artistiques. Il se caractérise par la très grande diversité des langues (plus de 100) et des aires culturelles concernées par ses activités de recherche et d'enseignement, ce qui lui confère une position académique unique et internationalement reconnue.

La numérisation des ressources (langues, textes, documentations, objets culturels) et leur traitement informatique constituent un enjeu majeur pour l'Inalco, à la fois en termes d'instruments de recherche (humanités numériques, accès aux données primaires, *Open Source Intelligence*, terrains numériques) et de valorisation des langues et des cultures (accès à la société de l'information des locuteurs de langues rares et informatiquement sous-dotées).

Au sein de l'Inalco, l'Équipe de Recherche Textes, Informatique, Multilinguisme (ERTIM, <https://ertim.inalco.fr>) est une unité de recherche spécialisée dans le Traitement Automatique des Langues (TAL) qui a pour objets de recherche le traitement des données multilingues, les

méthodologies du TAL multilingue et les applications en ingénierie linguistique. L'équipe exploite ou produit des méthodes et des outils pour l'analyse et le traitement des langues et des textes écrits ou oraux. Ses activités collaboratives relèvent soit des humanités numériques et de la description de langues spécifiques (notamment chinois, arabe, vietnamien, bambara, birman, quechua, hindi), soit d'applications ingénieriques (filtrage, classification, extraction d'information, etc.).

Pour renforcer ses activités, l'ERTIM recherche un ou une candidate à l'état de l'art en TAL (méthodes d'apprentissage profond, grands modèles de langues, traitement de la parole) apte à élaborer et développer des recherches pour l'analyse et le traitement des langues peu dotées. La personne recrutée doit faire valoir une expérience en recherche théorique et appliquée la rendant apte à aborder de manière innovante des problématiques complexes en termes de modélisation linguistique (modèle de langues sous-dotées en ressources ou ressources lacunaires) mais aussi discursives (gestion de l'alternance codique, créolisation, variation à la norme, contact de langues, etc.). Dans le cadre de cette CPJ et dans la perspective de la titularisation, la personne recrutée doit être en mesure de conceptualiser et superviser les tâches nécessaires en amont de la modélisation (traitement de l'oral, reconnaissance de la parole, transcription écrite, reconnaissance de l'écrit) et en aval (par exemple, traduction automatique).

Par son expertise et en lien avec ses projets de recherche, la personne recrutée pourra également être associée aux réflexions menées au sein de l'établissement sur l'usage de l'intelligence artificielle dans l'enseignement et la recherche en langues et en études aréales fondées sur des sources en langues orientales. Elle sera notamment impliquée dans l'élaboration d'une offre d'outils pédagogiques pour les enseignants de langue souhaitant intégrer les usages de l'IA, et d'outils méthodologiques destinés aux étudiants et chercheurs privés d'accès physique à leurs terrains de recherche.

Unités de rattachement : ERTIM

Lieu d'exercice : 2 rue de Lille (Paris)

Directeur de laboratoire : Damien Nouvel

N° de téléphone du directeur de laboratoire : 01 81 70 10 37

Courriel du directeur de laboratoire : damien.nouvel@inalco.fr

URL du laboratoire : <https://ertim.inalco.fr>

Projet d'enseignement

La personne recrutée sera amenée à dispenser des cours de 1^{er} et de 2^e cycle (Licence, Master) dans la formation TAL de l'Inalco et, de manière plus transversale, dans le cadre des formations aux humanités numériques et à la méthodologie du traitement de données numériques dispensées en master et en doctorat. Le public a pour spécificité d'être souvent issu de cursus en Sciences Humaines et Sociales et d'être composé de locuteurs ou d'apprenants de langues du domaine Inalco (Asie, Afrique, Amériques, Océanie, Europe orientale). Les cours seront choisis en concertation avec l'équipe pédagogique et en fonction des compétences propres de la personne recrutée. Ils peuvent être génériques (algorithmique, programmation, etc.), méthodologiques (traitement et contextualisation de données issues de terrains numériques) ou propres à ses domaines d'expertise. Une création de cours pourra être envisagée, en lien avec les objectifs pédagogiques des formations existantes.

Il est à noter que le Master TAL est co-accrédité avec les Université Sorbonne Nouvelle et l'Université Paris Nanterre. Il comprend des parcours recherches et professionnalisant débouchant sur des métiers de *data scientists*, ingénieurs NLP, ingénieurs linguistes, etc.

La charge d'enseignement est de 64h équivalent TD la première année, 96h équivalent TD les deuxième et troisième années, 128h équivalent TD en quatrième et cinquième année. A l'issue de la période de contractualisation, s'il est titularisé en tant que professeur des universités, le titulaire de la chaire délivrera un service réglementaire de 192 heures équivalent TD.

Département de rattachement : TIM

Lieu d'exercice : 2 rue de Lille (Paris) et 65 rue des Grands Moulins (Paris)

Directeur du département : Mathieu Valette

N° de téléphone du directeur du département:

Courriel du directeur du département : mvalette@inalco.fr

URL du département : <https://www.inalco.fr/textes-informatique-multilinguisme>

Condition du recrutement

Au terme des 5 années de pré-titularisation, la personne recrutée devra avoir fait la preuve qu'elle peut prétendre au titre de professeur des universités en termes de prise de responsabilités collectives, d'encadrement d'étudiants, de stagiaires, d'encadrement ou co-encadrement de doctorants et, enfin, de participation à la vie académique de l'équipe et de l'établissement.

La maîtrise du français n'est pas exigée au moment du recrutement mais il est attendu que le ou la candidate ait acquis un niveau suffisant à l'issue du contrat, soit au moment de la titularisation.

Modalités de candidature : Enregistrement des candidatures et dépôt du dossier de candidature au format PDF sur l'application ministérielle dédiée : <https://galaxie.enseignementsup-recherche.gouv.fr/antares/can/index.jsp>

Du 11/07/2024 à 10h jusqu'au 02/09/2024 à 16h, heure de Paris

Le dossier sera constitué de :

- **Formulaire de candidature à téléverser sur votre dossier de candidature Galaxie :** Le formulaire de candidature est disponible sur le site de l'Inalco : <https://www.inalco.fr/institut/concours-recrutement/chaire-de-professeur-junior-intelligence-artificielle-pour-les> (tout en bas de la page)
- Pièce d'identité avec photographie,
- Copie du diplôme de doctorat, ou d'un diplôme équivalent,
- CV analytique avec liste complète des publications
- 3 publications marquantes
- Présentation des projets de recherche, d'enseignement et d'insertion au sein de l'établissement et dans l'environnement de la chaire (10 pages maximum)
- Thèse de doctorat (et le rapport de soutenance le cas échéant) et résumé de la thèse en anglais
- Mémoire et dossier d'habilitation à diriger des recherches le cas échéant ou tout autre document justifiant d'une aptitude à l'encadrement doctoral.

Modalités de sélection et auditions :

L'évaluation sera réalisée par une commission de recrutement composée d'expert(e)s internes et externes. Seuls les candidats présélectionnés par la commission, à partir de l'examen des dossiers, seront convoqués à une audition.

Les critères d'évaluation porteront sur l'excellence du candidat, sa motivation, sa capacité d'encadrement ; sur la qualité et l'originalité des projets de recherche et d'enseignement ; sur sa capacité à intégrer son projet au sein du laboratoire, et à coordonner les activités du programme de la Chaire, ainsi que sur sa capacité à établir des réseaux collaboratifs à l'échelle internationale.

L'audition pourra se tenir en français ou en anglais.

CALL FOR APPLICATIONS

Junior Professor Chair:

Artificial Intelligence for Rare or Low-Resource Languages

JOB PROFILE

The National Institute of Oriental Languages and Civilizations invites applications for a Junior Professorship in "Artificial Intelligence for Rare or Low- Resource Languages".

This is a fixed-term contract (5 years) with a reduced teaching load, leading eventually, after evaluation, to tenure as a full professor.

Tenure begins no later than **December 31, 2024**.

The scheme takes place in two stages:

1. A call for applications for a research and teaching project, leading to the signing of a fixed-term contract as a public sector contract employee, for a period of 5 years;
2. At the end of the contract period, a tenure commission will meet to assess the candidate's scientific merit and professional aptitude, enabling them to be granted tenure in the corps of university professors. At the time of tenure, an HDR will be required. Publication of at least 6 articles in peer-reviewed journals is expected at the end of the 5-year pre-tenure period. During the pre-tenure period, the Chairholder is also expected to participate in collective projects initiated within the establishment, and to provide significant contributions for promoting the program (conferences, seminars, dissemination, etc.).

Contract for junior professorship: 5-year fixed-term contract under public law

CNU Section: 27, 07

Key words: Artificial Intelligence; Natural Language Processing; Low Resource Languages; Digitization; Speech Recognition

Salary: Index 735 of the French civil service (approx. 3,680€ gross per month)

Amount of associated funding: The Junior Professorship will receive a 200,000€ research grant from the French National Research Agency (ANR) for the 5-year period. The Chair program may

also benefit from additional funding from other national and international partners. This may include funding for project managers, contract PhD students, post-docs, conference participation, access to databases, equipment, subscriptions to services (algorithms or data), etc.

Candidate requirements: Candidates should hold a PhD, and have a well-established research record showing research activity in areas relevant to the job profile (see below), such as publications in leading international journals, scientific involvement in research projects, etc.

The successful candidate will need to demonstrate proficiency in one of the languages of their area of specialization, as well as the ability to teach in English.

Research project

The National Institute of Oriental Languages and Civilizations (Inalco) is a major public institute of higher education and research, renowned for its work on the languages, cultures and societies of the world through human and social science disciplines such as linguistics, sociology, economics, geography, history, international relations, political science, literary and artistic studies. Inalco is characterized by the great diversity of languages (over 100) and cultural areas covered by its research and teaching activities. This particularity makes Inalco a unique and internationally recognized academic institute.

The digitization of resources (languages, texts, documentations, cultural objects) and their analysis represent a major challenge for Inalco, both in terms of research tools (digital humanities, access to primary data, open-source intelligence, digital fields) and the promotion of languages and cultures (access to the information society for speakers of rare and low-resourced languages).

Within Inalco, the Texts, Computer Science, Multilingualism Research Team (ERTIM, <https://ertim.inalco.fr>) is a research unit specializing in Natural Language Processing (NLP), with a focus on multilingual data processing, multilingual NLP methodologies and their applications to linguistic engineering. The team uses or creates methods and tools for the analysis and processing of languages in written or spoken datasets. Its collaborative activities focus on Digital Humanities and the description of specific languages (notably Chinese, Arabic, Vietnamese, Bambara, Burmese, Quechua and Hindi), as well as engineering applications (text mining, text classification, text annotation, information extraction, etc.).

ERTIM is looking for a highly experienced candidate, mainly in the field of NLP (Deep Learning methods, Large Language Models, Speech Processing), capable of designing and developing research for the analysis and processing of low-resourced languages. The successful candidate will be able to demonstrate theoretical and applied research experience to tackle complex issues in innovative ways, both in terms of linguistic modeling (models of low-resourced or insufficiently resourced languages) and discursive modeling (managing code-switching, creolization, variation from the norm, language contact, etc.). As part of the chair and with a view to tenure, the person recruited must be able to conceptualize and supervise the tasks required

prior to modeling (speech processing and recognition, OCR) and following (e.g. automatic translation).

Through its expertise and in line with the research projects, the person recruited will also be expected to participate in discussions within the establishment on the use of Artificial Intelligence in teaching and research, for languages and areal studies based on oriental language sources. More specifically, the chair will be involved in the development of a range of teaching tools for language teachers wishing to integrate the use of AI, and methodological tools for students and researchers deprived of physical access to their research fields.

Affiliated research unit: ERTIM

Location: 2 rue de Lille (Paris)

Laboratory director: Damien Nouvel

Laboratory director's telephone number: 01 81 70 10 37

Laboratory director's e-mail: damien.nouvel@inalco.fr

Laboratory URL: <https://ertim.inalco.fr>

Teaching project

The person recruited will teach 1st and 2nd cycle courses (undergraduate, graduate) in Inalco's NLP training program and, more transversally, as part of the trainings in Digital Humanities and language data processing methodology, at master's and doctoral levels. The specificity of the audience is that they often come from Humanities and social sciences backgrounds and are speakers (or learners) of languages from the Inalco domain (Asia, Africa, the Americas, Oceania, Eastern Europe). Courses will be chosen in consultation with the teaching team and according to the skills of the person recruited. They may be generic (algorithms, programming, etc.), methodological (processing and contextualization of data from digital fields) or specific to the candidate's areas of expertise. The creation of new courses could be explored, in line with the pedagogical objectives of existing programs.

It should be noted that the NLP Master is co-accredited with Sorbonne Nouvelle University and Paris Nanterre University. It includes research and professional tracks leading to careers as data scientists, NLP engineers, linguist engineers, etc.

The teaching load is 64 hours in the first year, 96 hours in the second and third years, 128 hours in the fourth and fifth years (those hours are "equivalent TD"). At the end of the contractual period, if hired as a full professor, the chairholder will deliver a regulatory service of 192 hours.

Reporting department: TIM

Location: 2 rue de Lille (Paris) and 65 rue des Grands Moulins (Paris)

Department director: Mathieu Valette

Department director's telephone number:

Department director's e-mail: mvalette@inalco.fr

Department URL: <https://www.inalco.fr/textes-informatique-multilinguisme>

Recruitment conditions

At the end of the 5-year pre-tenure period, the person recruited must have demonstrated that they qualify for the title of university professor in terms of taking on collective responsibilities, supervising students and trainees, supervising or co-supervising doctoral students and, lastly, participating in the academic life of the team and the establishment.

Fluency in French is not required at the time of recruitment, but the candidate is expected to have acquired a sufficient level by the end of the contract, i.e. at the time of tenure.

Application procedures: Registration of applications and submission of the application file in PDF format on the dedicated ministerial application: <https://galaxie.enseignementsup-recherche.gouv.fr/antares/can/index.jsp>

From July 11th, 2024, at 10 a.m. to September 2th, 2024, at 4 p.m., Paris time

Application forms must include:

- **Application form to upload to your Galaxie application file:** The application form is available on the Inalco website <https://www.inalco.fr/institut/concours-recrutement/chaire-de-professeur-junior-intelligence-artificielle-pour-les> (at the very bottom of the page)
- Photo ID,
- Copy of doctoral degree or equivalent diploma,
- Analytical CV with full list of publications
- 3 outstanding publications
- Presentation of research, teaching and integration projects within the institution and the environment of the Chair (10 pages maximum)

- Doctoral dissertation (and defense report if applicable) and summary of dissertation in English
- Dissertation and (“Habilitation à Diriger des Recherches”), or any other document demonstrating aptitude for doctoral supervision.

Selection procedures and auditions:

Candidates will be evaluated by a recruitment committee comprising internal and external experts. Only those candidates shortlisted by the committee will be invited to an audition based on a review of their applications.

Evaluation criteria will focus on the candidate's excellence, motivation and supervisory skills; the quality and originality of their research and teaching projects; their ability to integrate their project within the laboratory, and to coordinate the Chair's program activities, as well as their ability to establish collaborative networks on an international scale.

The audition may be held in French or English.