



Offre n°2025-08903

PhD Position F/M 3-year PhD position in Computational Models of Semantic Memory

Type de contrat : Fixed-term contract

Niveau de diplôme exigé : Graduate degree or equivalent

Fonction : PhD Position

A propos du centre ou de la direction fonctionnelle

The Inria Lille - Nord Europe Research Centre was founded in 2008 and employs a staff of 360, including 300 scientists working in sixteen research teams. Recognised for its outstanding contribution to the socio-economic development of the Hauts-De-France région, the Inria Lille - Nord Europe Research Centre undertakes research in the field of computer science in collaboration with a range of academic, institutional and industrial partners.

The strategy of the Centre is to develop an internationally renowned centre of excellence with a significant impact on the City of Lille and its surrounding area. It works to achieve this by pursuing a range of ambitious research projects in such fields of computer science as the intelligence of data and adaptive software systems. Building on the synergies between research and industry, Inria is a major contributor to skills and technology transfer in the field of computer science.

Contexte et atouts du poste

The PhD position will be hosted within the MAGNET team at Inria Lille [1], in partnership with the SCALAB group at University of Lille [2] in an effort to strengthen collaborations between these two research teams, and specifically to foster cross-fertilizations

between Natural Language Processing (NLP) and psycholinguistics. The MAGNET is actually evolving into a new interdisciplinary research group focusing on cognitively-grounded computational, neural-based models of language and reasoning.

Mission confiée

This PhD project investigates semantic memory through complementary contrastive and integrative approaches, at the intersection of cognitive psychology and natural language processing. The overarching goal is to better understand the semantic capacities of large language models (LLMs) by comparing them to human cognition, and to improve these models using cognitively inspired learning biases.

Principales activités

The first research axis focuses on contrastive evaluation: we will design robust probing and prompting techniques to analyze how different families of LLMs (e.g., auto-regressive vs. masked models) encode and organize semantic knowledge. Models will be evaluated on datasets from experimental psychology, such as typicality norms (e.g., Rosch) and semantic feature norms (e.g., McRae, Buchanan), possibly including new data collection. The goal is to assess whether and how these models exhibit well-known features of human semantic memory such as taxonomic and prototypical organization, semantic feature sharing and inheritance, and polysemy —building upon preliminary work carried out in the team [3, 4, 5]. In addition, we intend to explore the structure of representations in vision-language models to investigate how multi-modal grounding shapes semantic memory, in light of findings from blind populations and developmental theories that challenge the necessity of visual input for acquiring rich word meanings.

The second axis focuses on integrative modeling, aiming to develop LLMs with inductive biases inspired by human cognitive development. Drawing from developmental psycholinguistics and findings in semantic memory acquisition, we will explore how representations evolve in humans and model this process in artificial learners. We will experiment with training regimes that control input volume, syntactic complexity, and curriculum structure. Longitudinal corpora and multimodal input (e.g., visual and symbolic data) will be used to simulate developmental conditions. This approach is directly inspired by recent initiatives such as the BabyLM benchmark campaigns, which promote the design of smaller, more data-efficient language models grounded in child language learning. Our goal is to integrate such developmental constraints into the architecture and training of LLMs in order to foster interpretability, efficiency, and cognitive plausibility. In both axes, both English and French data will be considered.

- [3] <https://aclanthology.org/2023.eacl-main.167.pdf>
[4] <https://aclanthology.org/2023.findings-emnlp.615.pdf>
[5] <https://aclanthology.org/2024.emnlp-main.156.pdf>

Avantages

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Rémunération

2 200€ Gross monthly salary (before taxes)

Informations générales

- **Thème/Domaine** : Data and Knowledge Representation and Processing Statistics (Big data) (BAP E)
- **Ville** : Villeneuve d'Ascq
- **Centre Inria** : [Centre Inria de l'Université de Lille](#)
- **Date de prise de fonction souhaitée** : 2025-10-01
- **Durée de contrat** : 3 years
- **Date limite pour postuler** : 2025-06-30

Contacts

- **Équipe Inria** : [MAGNET](#)
- **Directeur de thèse** :
Denis Pascal / Pascal.Denis@inria.fr

A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

L'essentiel pour réussir

Applicants should hold a Master's degree (or equivalent) in one or more of the following fields: Computational Linguistics, Natural Language Processing, Artificial Intelligence, Machine Learning, Cognitive Science. Strong programming skills (Python preferred), a solid foundation in empirical research methods, and an interest in interdisciplinary work combining formal, computational, and experimental approaches are highly desirable.

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

Consignes pour postuler

Applications will be considered until the position is filled. However, you are encouraged to apply early as we shall start processing the applications as and when they are received.

Applications, written in English, should be submitted online and should include:

- * Curriculum Vitae (including your contact address, work experience, publications)
- * Cover letter indicating your research interests and your motivation
- * Contact information for at least 2 referees

Applications should be sent to: Pascal Denis (pascal.denis@inria.fr) et Angèle Brunellière (angele.brunelliere@univ-lille.fr).

Sécurité défense :

Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement :

Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.