



**Offre n°2022-05389**

## **PhD Position F/M 3-year PhD position in Automatic Argumentation Mining in French Legal Decisions**

*Le descriptif de l'offre ci-dessous est en Anglais*

**Type de contrat :** CDD

**Niveau de diplôme exigé :** Bac + 5 ou équivalent

**Fonction :** Doctorant

### **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**

We invite applications for a 3-year PhD position co-funded by Inria, the French national research institute in Computer Science and Applied Mathematics, and LexisNexis France, leader of legal information in France and subsidiary of the RELX Group.

The position is affiliated with the MAGNET, a research group at Inria, Lille, which has expertise in Machine Learning and Natural Language Processing, in particular Discourse Processing. The PhD student will also work in close collaboration with the R&D team at LexisNexis France, who will provide their expertise in the legal domain and the data they have collected.

### **Mission confiée**

The overall objective of this project is to develop an automated system for detecting argumentation structures in French legal decisions, using recent machine learning-based approaches (i.e. deep learning approaches). In the general case, these structures take the form of a directed labeled graph, whose nodes are the elements of the text (propositions or groups of propositions, not necessarily contiguous) which serve as components of the argument, and edges are relations that signal the argumentative connections between them (e.g., support, offensive). By revealing the argumentation structure behind legal decisions, such a system will provide a crucial milestone towards their detailed understanding, their use by legal professionals, and above all contributes to greater transparency of justice.

### **Principales activités**

The main challenges of this project start with the creation and release of a large-scale dataset of French legal decisions annotated with argumentation structures. To minimize the manual annotation effort, we will resort to semi-supervised and transfer learning techniques to leverage existing argument mining corpora, such as the European Court of Human Rights (ECHR) corpus, as well as annotations already started by LexisNexis. Another promising research direction, which is likely to improve over state-of-the-art approaches, is to better model the dependencies between the different sub-tasks (argument span detection, argument typing, etc.) instead of learning these tasks independently. A third research

avenue is to find innovative ways to inject the domain knowledge (in particular the rich legal ontology developed by LexisNexis) to enrich the representations used in these models. Finally, we would like to take advantage of other discourse structures, such as coreference and rhetorical relations, conceived as auxiliary tasks in a multi-tasking architecture.

## Compétences

The successful candidate holds a Master's degree in computational linguistics, natural language processing, machine learning, ideally with prior experience in legal document processing and discourse processing.

Furthermore, the candidate will provide strong programming skills, expertise in machine learning approaches and is eager to work at the interplay between academia and industry.

## 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

1st and 2nd year : 2051 € Gross monthly salary (before taxes)

3rd year : 2158 € gross monthly salary (before taxes)

## Informations générales

- **Thème/Domaine** : Représentation et traitement des données et des connaissances Statistiques (Big data) (BAP E)
- **Ville** : Villeneuve d'Ascq
- **Centre Inria** : [Centre Inria de l'Université de Lille](#)
- **Date de prise de fonction souhaitée** : 2023-03-01
- **Durée de contrat** : 3 ans
- **Date limite pour postuler** : 2023-02-16

## Contacts

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

The successful candidate holds a Master's degree in computational linguistics, natural language processing, machine learning, ideally with prior experience in legal document processing and discourse processing. Furthermore, the candidate will provide strong programming skills, expertise in machine learning approaches and is eager to work at the interplay between academia and industry.

**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: Aurélien Bellet, Pascal Denis, and  
Mikaella Keller (firstname.lastname@inria.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.