

Le projet ANR CHAPITRES recrute, à partir de septembre 2016, une ingénieure ou un ingénieur de recherche ou d'études à temps plein.

**Une ingénieure ou un ingénieur de recherche ou d'études à temps plein**

Durée : un an, renouvelable pour un an

Estimation du salaire mensuel net : 1861 € (en fonction des qualifications)

### Missions

- aide à la conception et construction de la base de données
- veille scientifique
- alimentation du carnet de recherches
- communication et diffusion

### Profil requis :

- connaissance du champ de la littérature française et de l'histoire littéraire des 19<sup>e</sup> et 20<sup>e</sup> siècles
- formation ou expérience en édition électronique
- maîtrise de la TEI et en particulier des métadonnées
- expérience des bases de données bibliographiques et des archives existantes telles que Gallica ou archive.org
- maîtrise d'un ou plusieurs logiciels de gestion de bases de données
- niveau Master ou doctorat

### Modalités de recrutement :

Le dossier de candidature, constitué d'un CV, d'une lettre de motivation et de deux articles parus ou en cours de publication, devra être envoyé au plus tard lundi 20 juin 2016, par e-mail uniquement, à [aude.leblond@gmail.com](mailto:aude.leblond@gmail.com). Les candidats recevront un message confirmant la bonne réception de leur dossier. Les candidats retenus à l'examen des dossiers seront convoqués pour un entretien qui aura lieu fin juin 2016.

Les candidats recrutés seront employés par l'université Sorbonne nouvelle – Paris 3 et rattachés au laboratoire THALIM, Théorie et Histoire des Arts et des Littératures de la Modernité (UMR 7172).

Les candidats sont invités à prendre connaissance des orientations scientifiques du projet Chapitres (voir annexe), afin de pouvoir présenter lors de l'audition la manière dont ils envisagent la réalisation des tâches techniques requises.

## Annexe : besoins techniques de l'équipe

L'ingénieur d'études ou de recherche sera dans un premier temps chargé de recueillir des données numérisées pré-existantes pour alimenter une base de données massive. Il s'agit d'aspirer un grand nombre de textes déjà en accès libre sur Internet, par exemple à partir de Gallica ou du Projet Gutenberg, avec leurs méta-données. Cela supposerait une expertise sur le choix de l'édition à prendre, le choix des textes à aspirer, qui pourra se faire en concertation avec l'équipe. Il faudrait ensuite vérifier la qualité des méta-données, veiller au balisage des divisions textuelles, en particulier les chapitres, et ajouter quelques informations supplémentaires sous forme de tags. On vise la constitution d'un corpus de quelques milliers de textes de prose narrative.

Ce corpus massif devra faire une place au canon littéraire, mais également à des textes non-canoniques : le choix de ces textes constitue donc un enjeu méthodologique important, auquel participera l'ingénieur d'études ou de recherche chargé de la constitution du corpus. Après un premier temps consacré au rassemblement d'un corpus canonique, il faudra dans un deuxième temps l'enrichir de textes non-canoniques. Ce deuxième moment supposera un travail plus approfondi sur l'histoire littéraire et pourrait en priorité intéresser les candidats ayant une formation en édition électronique.

En coordination avec les chercheurs de l'équipe, on pourra dans un deuxième temps passer à une phase de data mining, qui permettra de comparer les longueurs des chapitres entre eux, l'évolution de cette longueur et des ratios par rapport au texte entier au fil du temps, mais aussi de mettre en évidence l'évolution historique des pratiques d'intitulation, de révéler d'éventuelles différences en fonction des sous-genres romanesques, etc. Il serait également intéressant de soumettre les débuts ou fins de chapitres à des requêtes lexicales, pour voir si ces seuils sont caractérisés par un lexique reconnaissable.

La deuxième année de travail, ainsi, pourra être enrichie par une participation accrue aux événements scientifiques prévus dans le projet (1 journée d'études et 3 colloques en plus d'un séminaire bimensuel), ainsi qu'à la publication des travaux de l'équipe, en ligne ou sur papier.