

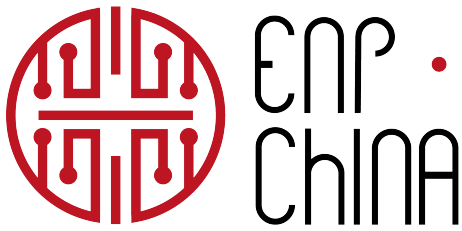
**Natural Language Processing PhD position in the
ERC Advanced Grant Project “Elites, networks, and
power in modern urban China (1830-1949)”
[ENP- CHINA]**

The Institute of Asian Studies (IrAsia), Laboratoire Parole et Langage, and Laboratoire d’Informatique et Systèmes at Aix-Marseille University invite applications for one doctoral fellowship in NLP in the ERC-funded project ENP-CHINA.

Events are crucial elements for historical analysis. Natural Language Processing (NLP) has been interested in events (Passoneau, 1988) and temporal structure (Mani et al, 2006) because of their importance for text understanding and for their necessity to model a wide range of linguistic phenomena. The notion of ‘Event’ in NLP comes from the one in linguistics where it had been heavily studied (Klein, 1994). In these fields ‘events’ are strongly associated with their linguistic expressions and to the precise temporal structure expressed by a sentence and ultimately a text, taking into accounts tensed verbs, temporal adverbs among many other things... However, events as defined in NLP differ from what constitute an event for historians. An historical event often refers to a peak in a continuum of actions that coalesce into a key event with a more or less enduring and precise denomination (e.g. May 4 Movement, Anti-British Boycott, etc.). These actions and events exist only through the varying and changing terms that characterize them in our case in press (news?) corpora.

The general objective of the PhD is to devise ways to mobilize NLP techniques (in particular the most recent advances in machine learning) in the detection of historical events and their related actors.

The successful candidate will have the task to bridge historical and



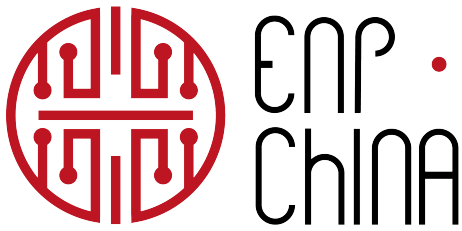
linguistic events by studying how event detection in NLP can help to detect and model historical events in the making. He will develop methods and tools to cluster news articles around key events based on outputs from deep NLP analysis of large press corpora and a database which describes historical actors.

Although the system will be developed and evaluated based on expert knowledge and well identified events, which typically received a name, the work will eventually lead to the detection and description of other historical events which may have gone unnoticed and address the issue of scale in event definition.

The research of the successful candidate will thus involve formalizing knowledge about events, improving and applying NLP tools and experimenting with unsupervised or weakly supervised machine learning algorithms. The candidate is expected to produce usable prototypes to be able to receive feedback from the historians in the project.

Context

The ENP-CHINA project proposes a step-change in the study of modern China reliant upon scalable data-rich history based on an array of cutting-edge digital methods within an integrated virtual research environment. The project envisions elites in urban China as actors whose status, position, and practices were shaped by the power configurations that developed over time and whose actions through institutions and informal/formal networks in turn were a determining factor in redrawing social and political boundaries. The key issue that the project will address is breaking through existing limits of access to historical information that is embedded in complex sources and its transformation into refined, re-usable and sustainable data for contemporary and future study of modern China.



Job description

As part of the project you will:

- Join an innovative research project funded by the prestigious ERC grant scheme
- Become a member of a large and fully interdisciplinary research team, including historians, linguists, computer and data scientists.
- Benefit from a vibrant internationally oriented research environment
- Attend international conferences, workshops, summer schools, and for research visits at institutions in the project countries

Tasks

- Participate in the development of methods and tools to process and analyze large scale press corpora,
- Develop one's own PhD project within the scope of the ENP-CHINA project and write (choice of English or French language) and complete the PhD dissertation within three years;
- Fully participate in the project activities and publications in collaboration with the other team members;
- Participate in international conferences, workshops, seminars and other scholarly activities.

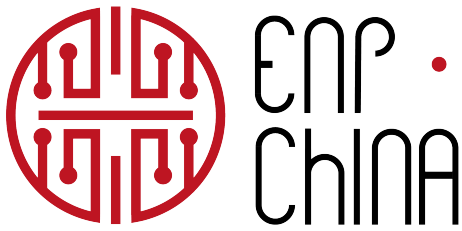
Requirements

Candidates should have the following credentials:

- A completed Research Master in Computer Science, NLP or applied mathematics, with strong programming skills and a lively interest in text processing.
- Familiarity with existing NLP toolkits and machine learning algorithms and libraries.
- Skills in web development to present and share interactive results (any language/framework)
- Ability, willingness and commitment to work in a multi-disciplinary team;
- Good command spoken and written English;

The application must include:

- CV including degrees and other completed courses, work experience and a list of degree projects/theses, language skills
- A letter (max. 1000 words, excl. references) describing your research interests, skills and motivation to apply for the position:
 - why you are interested in the field/project described in the announcement
 - what makes you suitable for the project in question
- An attested copy of relevant degree(s)
- Copies of elements that illustrate one's abilities (theses, articles etc.)
- Two reference letters with contact details [The letters of reference must be sent by the referees directly to Prof. Christian Henriot at enpmuc@gmail.com.]



Conditions of employment

The PhD research position is for 3 years. The appointment will be initially for 12 months. After satisfactory evaluation of the initial appointment, it can be extended for a total duration of three years.

Starting date: 1 September 2019

Registration in one of Aix-Marseille University's PhD programs (Department of Computing or Department of Language Sciences)

Salary scale (net starting salary) is 1,400 Euros per month before taxes. The package includes full social benefits, complementary insurance (mutuelle), and pre-paid meal tickets (ticket restaurant). Candidates will be provided with new computer, office space, and benefit from research allowances for field trip/conference (subject to P.I.'s approval).

General French regulations and dispositions apply for pension scheme, work time, holidays, and all social security issues

Regular presence in the offices of the team and participation to work, brainstorming, and academic meetings

Selection

The selection among the eligible candidates will be based on previous experience and grades, the quality of the degree thesis, references, relevant experience, and the candidate's written motivation for seeking the position. Key selection criteria for the doctoral positions are analytical skills, methodological abilities, language proficiency and the capacity for team-work as well as independent research. After a careful examination of the applications, the short-listed candidates (no more than 3 for one position) will be interviewed.



How to apply

Please send your application in one PDF document to enpmuc@gmail.com. Use the subject heading “ENP-CHINA NLP PhD Position” .

Deadline: The position will remain open until a suitable candidate is selected. We shall process applications as they arrive with the first round of application review on 31 July 2019.

If you have any further inquiries about the position, please contact Dr. Pierre Magistry at pierre.magistry@univ-amu.fr



This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 788476).