Online social platforms provide information about the interaction structure of users forming globally connected networks, where information is exchanged commonly in textual form. While users can be associated to their individual socio-demographic status (SES), linguistic features, or topical interest, they may form well connected communities around commonly interesting subjects. This coupling associated to their individual socio-demographic status (SES), linguistic features, or topical interest, may form well connected communities around commonly interesting subjects. This coupling appears to be a fundamental character of online social networks. However, current approaches usually consider only one side of these connections...
correlations, while fail to identify multivariate dependencies between the network, SES, and language. To reach this integration, we aim to take advantage of the recently developed methodological framework of deep neural networks.

The objectives of this postdoc is to explore this context via the development of semi-supervised multi-factorial analysis methods based on deep learning using heterogeneous data. The aim is to infer correlations / patterns that exist between dynamic linguistic variables, mesoscopic structure and social network dynamics and their socio-demographic and socio-economic attributes. We will base our studies on a corpus of 200 Million tweets from 2 Million users that has been collected over a period of 2 years within the DANTE team. In addition to this corpus itself, we combine this textual data with network informations inferred from tweets and user profiles, and socioeconomic informations obtained from census data.

Main activities
The successful applicant will work on the following topics:

- Collection, statistical analysis, and mining of large digital datasets recording the blog posting activity and interactions of millions of users on Twitter.
- To development of statistical learning methods to understand the co-evolution of the network structure, language dynamics, and individual variables like SES, location, or demographic variables.
- To carry innovative methods for the representation of results.

Skills
Applicants should have a PhD degree in computer science, physics, or related discipline with strong interest in complex networks, social phenomena, and computational linguistic.

Background in complex networks, data analysis, computational modelling is an advantage.

Efficiency in programming, data collection and analysis are required.

Good academic writing and presentation skills in English are required.

There are no teaching obligations but opportunities.

Benefits package

- Subsidised catering service
- Partially-reimbursed public transport
- Social security
- Paid leave
- Flexible working hours
- Sports facilities

Remuneration

Gross salary: 2650 Euros per month