datasets for use in predictive data-driven models. Our main novelty is to define network reduction for modelling social contagion processes. A propos du centre ou de la direction fonctionnelle
Grenoble Rhône-Alpes Research Center groups together a few less than 800 people in 35 research teams and 9 research support departments. Staff is located on 5 campuses in Grenoble and Lyon, in close collaboration with labs, research and higher education institutions in Grenoble and Lyon, but also with the economic players in these areas. Present in the fields of software, high-performance computing, Internet of Things, image and data, but also simulation in oceanography and biology, it participate at the best level of international scientific achievements and collaborations in both Europe and the rest of the world.

Contexte et atouts du poste
Within the framework of a partnership
• public with French National Research Agency (ANR) : DATAREDUX, a partnership between:
  • Dante at LIPI, ENS de Lyon - Paulo Gonçalves (Inria)
  • Centre de Physique Théorique - Alain Barrat (CNRS)
  • Institut Pierre Louis d'épidémiologie et de santé publique - Vittoria Colizza (INSERM)
  • External collaboration with Márton Karsai (ECU Vienna, Austria)

Founded in 1880, the Ecole Normale Supérieure de Lyon (www.ens-lyon.fr) is one of the most prestigious "Grande École" in France. It selects students from the top 5% among all students in the country and among the best students from abroad. The mission of the university is training national and international future professors, researchers, senior civil servants as well as business and political leaders, and to advance research in several disciplines ranging from humanity to natural sciences.

Established in 1967, Inria (www.inria.fr) is the only public research body in France fully dedicated to computational sciences. Combining computer sciences with mathematics, Inria's 3,500 researchers strive to invent the digital technologies of the future. Educated at leading international universities, they creatively integrate basic research with applied research and dedicate themselves to solving real problems, collaborating with the main players in public and private research in France and abroad and transferring the fruits of their work to innovative companies.

The DANTE (team.inria.fr/dante/) team is and Inria team located in Lyon hosted by the ENS Lyon and the IXXI Complex System Institute. The goal of DANTE is to study and model the dynamics of large-scale complex networks, e.g. social networks, technological networks, social-communication networks, etc. Building on the advancements of the digital data revolution, our main challenge is to propose generic methodologies and concepts to develop relevant formal tools to model and analyse the dynamics of dynamical networks and ongoing dynamical phenomena. Our main focus areas are:
• Access and collect data with adapted and efficient tools. This includes a reflexive step on the biases of the data collected and their relations to real activities/application domain.
• Characterise and model the dynamics of complex temporal network structures (a) through the development of novel representation and metrics, (b) through the analysis of sizeable real-world datasets; and (c) through large-scale numerical simulation based on novel computational models. We are interested in the structural and temporal properties of dynamical networks and their evolution with ongoing dynamical processes.
• Invent original approaches combining graph theory with signal processing. A key point is to capture temporal features in the data, which may reveal meaningful insights on the evolution of the networks.
• Find solutions in the applications domain related to communication networks. In particular, we study the dynamic of the network (available bandwidth, traffic, etc.) and propose solutions to adapt its parameterisation accordingly.

Mission confiée
Many rich datasets on actions and interactions of individuals have recently become available, commonly encoded as networked systems, arising from heterogeneous sources with details at different scales and resolutions, and potentially containing geographical and temporal information as well as metadata. These outstanding sources of information and knowledge fuel a wide spectrum of data-driven numerical simulations of dynamical processes. Data alone, however, even in huge amounts, do not easily transform into knowledge or predictive models. The rich and diverse information they contain raise crucial challenges concerning their analysis, representation and interpretation, the extraction of meaningful structures, and their integration into data-driven models.

In this context, this PhD program as part of the DataRedux project aims to contribute to the development of an innovative framework to reduce networked data complexity while preserving its richness, by working at intermediate scales ("mesoscales"). Our objective is to reach a fundamental breakthrough in the theoretical understanding and representation of rich and complex networked datasets for use in predictive data-driven models. Our main novelty is to define network reduction for modelling social contagion processes.
techniques in relation with the dynamical processes occurring on the networks. To this aim, we will develop methods to go from data to information and knowledge at different scales in a human-accessible way by extracting structures from high-resolution, diverse and heterogeneous data. Our methodology will involve the identification of the most relevant subparts of time-resolved datasets while remapping the remaining parts of the system, the simultaneous structural-temporal representations of time-varying networks, the development of parsimonious data representations extracting meaningful structures at mesoscales, and the building of models of interactions that include mesostructures of various types. Our aim is to identify data aggregation methods at intermediate scales and new types of data representations that carry the richness of information of the original data, keeping their most relevant patterns and summarising less salient properties for their more manageable integration in data-driven models for decision making and actionable insights.

Principales activités
During the program the successful candidate will analyse, and model temporal networks and dynamical processes intermediated by them. Special focus will be payed for the modelling of social contagion phenomena and information diffusion. To achieve these goals the candidate is going to carry out research in the following directions:

- Observing, analysing, and modelling temporal patterns of interactions to develop new network reduction and filtering techniques using methods borrowed from statistical network analysis, data mining, and representation learning to automatically extract structures from data, creating filters that focus progressively on more salient parts of a dataset and to define new data representations that preserve their essential properties while summarizing unnecessary details.

- Application of the developed methods on large-scale real temporal network datasets to assess the impact of incomplete or noisy data on the data reduction methods and to identify ways how to deal with such imperfections in data-driven models of dynamical processes.

- Validation of the newly developed methods with respect to their use in data-driven models of contagion phenomena. Goal is to explore the relevance of the extracted structures in case of social contagion phenomena with the aim to distinguish between simple and complex contagion processes in models and in empirical settings.

Among other possible techniques to achieve these objectives, a particular focus is envisioned on signal processing and machine learning approaches applied on networks revolving around:

- Metric learning, to apprehend the most relevant notion of distances / homophily between agents (human centric Machine Learning), with respect to the referred task;
- Graph learning under sparsity constraints (e.g. graphical lasso);
- Data dimension reduction via graph coarsening as well as via sketching with random feature maps.

This data-driven program will strongly rely on large corpuses of human dynamical dataset enriched with precise information about the underlying social network. The candidate will develop computational methods at the crossing of computational social science, data science, computer science, and network science. During the training the candidate will optimally benefit from the diverse expertise of the supervisors and the hosting and collaborating teams. Common work will be maintained with the Statistical Physics and Complex Systems team from CNRS, CPT Marseille and the Pierre Louis Institute of Epidemiology and Public Health from INSERM. The candidate will be hosted by the Inria DANTE team at LIP ENS Lyon. The program will be directed by Dr. Paulo Goncalves (Research Director, Inria) and co-directed by Dr. Márton Karsai (Associate Professor, Department of Network and Data Science, Central European University) with the potential involvement of Dr. Rémi Gribonval (Research Director, Inria) and Pierre Borgnat (Research Director, CNRS).

Compétences
Applicants at PhD level should have a master's degree in computer science, physics, data science or related discipline with strong interest in studying complex networks and human dynamics. Background in at least one of the following areas: network science, data science, statistical signal processing, machine learning, computational modelling, is expected, proven by excellent grades.

Efficiency in programming, data collection, and analysis are required together with good academic writing and presentation skills in English.

There are no teaching obligations but opportunities.

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 (after 6 months of employment) 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: 1 982 euros brut /month
3rd year: 2 085 euros brut / month