Call for Visiting Professors

DATAIA Institute

General context of the call

DATAIA is the University of Paris-Saclay's artificial intelligence institute dedicated to data science, AI and their disciplinary and application interfaces. It brings together over 800 professors and researchers from universities, national research organizations and Grandes Écoles. Funded as a convergence institute by the Agence Nationale de la Recherche (ANR) as part of the "Programmes d'Investissements d'Avenir", DATAIA aims to :

- Mobilize multidisciplinary skills to produce new knowledge;
- Develop innovative training programs of excellence at Master's and PhD levels, for both initial and continuing education, based on this interdisciplinary research;
- Structure long-term international and industrial collaborations at the highest scientific level around the main challenges addressed by the Institute (https://www.dataia.eu/en/linstitut-dataia-paris-saclay).

Scientific scope of the call and objectives

Responses to this call must show how they contribute to the objectives of the DATAIA Institute, and in particular to the structuring of the data science field within the Paris-Saclay Campus. Consequently, all proposals must fall within the scope of at least one of DATAIA's 4 main challenges, and may also concern an application area (see the list of "Axes de travail" in Appendix 1). Guests must be attached to at least two research teams belonging to DATAIA Institute partner institutions (see list in Appendix 2).

This call for proposals will finance invitations, on an ad hoc basis, for 3 to 12 months, with no start date constraints. The invited person must hold a research or teaching position in a foreign university or research organization. DATAIA finances the visit with €4,000 per month in the form of a stipend, to be used for the visitor's remuneration and/or living expenses. DATAIA provides an additional allowance for travel expenses:

- Up to **€1,000** for a guest from a European country;
- Up to **€1,500** for a guest from outside Europe

Expectations/Commitments

The visitor undertakes to offer a cycle of courses or seminars at Master's or PhD level in the partner laboratories. For long-term stays, the jury will assess the visitor's commitment to student co-supervision (e.g., M2 internships, PhDs, post-docs, etc.) and to setting up collaborations.

A one- to two-page summary of the visitor's activities will be drawn up by one of the host laboratories within 1 month of the visit. This report will detail activities carried out during the visit, highlights and prospects for future collaborations. The guest also undertakes to report to the DATAIA Institute on any facts that enable the impact of the visit to be measured (joint projects or publications in particular).

Recipients of DATAIA Institute support must acknowledge DATAIA and the Programmes d'Investissement d'Avenir (PIA) in all scientific communications associated with funded activities by inserting:

"This research was supported by DATAIA convergence institute as part of the "Programme d'Investissement d'Avenir", (ANR-17-CONV-0003) operated by [Partner] XXX".

Partnership scope of the call

This call for proposals is reserved for teams within the scope of the DATAIA Institute. Only the laboratories of partner establishments listed in the funding agreement will be able to carry and manage the allocated credits (see list in Appendix 2).

How to submit a proposal?

The dossier will be drawn up jointly by the host laboratories and the person being considered for an invitation. The application must be submitted by one of the host laboratories by email to: submission-dataia@inria.fr. **Questions about this call can be sent to: contact-dataia@inria.fr.**

The application must include administrative and financial information, a description of the scientific file and the following elements:

- Candidate's CV (3-5 pages, including contact details, status and home institution, as well as main scientific contributions);
- Scientific focus of DATAIA Institute (cf. Appendix 1);
- Program of work during the visit (1 page max);
- Proposed courses / seminars (1-2 page(s) max);
- Expected benefits for the DATAIA Institute.

Project selection procedure

Projects will be reviewed and selected by DATAIA's Executive Committee and Program Committee, and a response will be given within one month of submission.

Appendix 1: DATAIA research areas

The scientific program of each proposal must explicitly address at least one of the four research areas outlined below, which are DATAIA's four interdisciplinary challenges.

FROM DATA TO KNOWLEDGE, FROM DATA TO DECISIONS

The growing availability of massive data is pushing back the technical frontiers in many fields. On the one hand, the heterogeneous, semi-structured, incomplete or uncertain nature of data calls into question the usual statistical models and algorithms dedicated to decision-making. On the other hand, data management raises new operability constraints, such as security, integrity and traceability. What's more, producing knowledge requires building models that deliver explainable, statistically valid and calculable decisions. Acceptance of results also requires that confidentiality and loyalty be reinforced. At the same time, new developments in optimization should make it possible to improve estimation procedures.

Challenges:

- Heterogeneous, complex, incomplete, semi-structured and/or unsure data
- Massive data: algorithms and data structuring
- E-learning, methodology for massive data, efficient methods
- Improved storage, computation and estimation for data science
- Game-theoretic modeling of interactions between agents (human or artificial)
- Multiscale and multimodal representation and algorithms
- Theoretical analysis of heuristic methods (complexity theory, information geometry, Markov chain theory)
- Human-machine coevolution in autonomous systems: conversational agents, cars, social robots.

LEARNING AND ARTIFICIAL INTELLIGENCE

Recently, deep learning research has made spectacular advances in computer vision and natural language processing. Beyond the arrival of massive data, increased computing power and design efforts, the causes of these advances, which are still poorly understood, raise at least three questions. What learning theory will enable us to analyze deep architectures? How can we manage the compositionality of these architectures and their ability to apprehend more complex objects? How can we open the black box to update learned representations?

Challenges:

- Innovative machine learning and AI: common sense, adaptability, Generalization
- Deep learning and adversarial learning
- Machine learning and hyper-optimization
- Optimization for learning (e.g. improvements in stochastic gradient methods, Bayesian optimization), combinatorial optimization
- Learning-modeling link, a priori integration in learning
- Reproducibility and robust learning
- Statistical inference and validation
- Compositionality of deep architectures.

TRANSPARENCY, RESPONSIBLE IA AND ETHICS

Digital trust is built on the implementation of ethically responsible methodologies through the transparency and accountability of algorithmic systems; the regulation of the collection, use and processing of personal data; and the reinforcement of regulation through appropriate digital procedures. Privacy by design is a form of regulation that includes the protection of personal data at all stages of collection and processing. The tracing of tools applied to data must also be developed in such a way as to facilitate the explanation of the model for experts and users alike, making algorithmic systems auditable. Confidentiality principles, although easy to formulate, require modifications to storage and processing infrastructures, with major legislative, sociological and economic impacts. Transparency techniques for algorithmic systems will be developed, focusing on: fairness, loyalty and non-discrimination, and accountability-by-construction.

Challenges:

- Accountability and explicability by design
- Transparency and fairness by design
- Auditing algorithmic systems: non-discrimination, fairness, technical bias, neutrality, equity
- Measuring trust and digital appropriation
- Progressive user-analysis of progressive data (interactive monitoring of decision-making systems)
- Responsibility for information processing and decision-making: data usage control and fact-checking
- Causal discovery, traceability of inferences from source data, interpretability of deep architectures

PROTECTION, REGULATION AND THE DATA ECONOMY

Companies involved in the data economy continually need to rethink how they are structured: they must adopt a project-oriented organization with rapid changes in resource allocation. The data economy also raises issues of concentration and monopoly. A small number of companies (GAFAM) hold most of the data. This market concentration can lead to unfair competition, and innovation in small and medium-sized businesses is likely to suffer. Citizens expect governments to intervene in the digital economy to prevent too much concentration and monopoly. Governments must prevent information leakage to preserve state sovereignty and respect for regulations.

Challenges:

- Privacy by GDPR design
- Privacy-aware learning (differential privacy)
- Development of ethically responsible methodologies and technologies to regulate the collection, use and processing of personal data, and the exploitation of knowledge derived from such data
- IT security for data processing chains
- Security/crypto: blockchain and trusted third parties

Projects may address topics more specific to particular fields of application. A list of these fields is given below as an indication. However, it is requested that projects address data science issues of general interest, i.e. not specifically restricted to one application area.

- Energy: optimization of management (production and distribution), and regulation
- Health, personalized nutrition and well-being
- Urbanization Mobility (connected and autonomous vehicles, smart cities, etc.)
- Analysis for finance and insurance
- Internet of people and things
- E-Sciences

Appendix 2: List of DATAIA partners eligible for the call for proposals

• Agro ParisTech • CentraleSupélec • CEA • CNRS • ENS Paris-Saclay • IFP- Énergies Nouvelles • Inria	• Institut Mines-Télécom Business School • Inrae • ONERA • Université d'Evry Val d'Essonne • Université Paris-Saclay • Université Versailles St-Quentin-en-Yvelines